Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **452,** 1502–1518 (2015) doi:10.1093/mnras/stv1416

Black hole evolution – I. Supernova-regulated black hole growth

Yohan Dubois,[1,2]★ Marta Volonteri,[1,2] Joseph Silk,[1,2,3,4] Julien Devriendt,[3,5] Adrianne Slyz[3] and Romain Teyssier[6]

[1]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[2]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[3]*Sub-department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, UK*
[4]*Department of Physics and Astronomy, The Johns Hopkins University Homewood Campus, Baltimore, MD 21218, USA*
[5]*Observatoire de Lyon, UMR 5574, 9 avenue Charles André, F-69561 Saint Genis Laval, France*
[6]*Institute für Theoretische Physik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*

Accepted 2015 June 23. Received 2015 June 15; in original form 2015 March 26

ABSTRACT

The growth of a supermassive black hole (BH) is determined by how much gas the host galaxy is able to feed it, which in turn is controlled by the cosmic environment, through galaxy mergers and accretion of cosmic flows that time how galaxies obtain their gas, and also by internal processes in the galaxy, such as star formation and feedback from stars and the BH itself. In this paper, we study the growth of a $10^{12}\,M_{\odot}$ halo at $z = 2$, which is the progenitor of a group of galaxies at $z = 0$, and of its central BH by means of a high-resolution zoomed cosmological simulation, the Seth simulation. We study the evolution of the BH driven by the accretion of cold gas in the galaxy, and explore the efficiency of the feedback from supernovae (SNe). For a relatively inefficient energy input from SNe, the BH grows at the Eddington rate from early times, and reaches self-regulation once it is massive enough. We find that at early cosmic times $z > 3.5$, efficient feedback from SNe forbids the formation of a settled disc as well as the accumulation of dense cold gas in the vicinity of the BH and starves the central compact object. As the galaxy and its halo accumulate mass, they become able to confine the nuclear inflows provided by major mergers and the BH grows at a sustained near-to-Eddington accretion rate. We argue that this mechanism should be ubiquitous amongst low-mass galaxies, corresponding to galaxies with a stellar mass below $\lesssim 10^{9}\,M_{\odot}$ in our simulations.

Key words: methods: numerical – galaxies: active – galaxies: evolution – galaxies: formation.

1 INTRODUCTION

Supermassive black holes (BH) are common compact objects observed in the centre of galaxies. They are suspected to grow along with their host galaxy as observations suggest a strong scaling between BH masses and galaxy properties (e.g. Magorrian et al. 1998; Tremaine et al. 2002; Häring & Rix 2004; Kormendy & Ho 2013). Gas accretion drives most of the mass growth of black holes at high-redshift, and this accretion mechanism returns a fraction of the rest-mass accreted energy into effective feedback for the host-galaxy, which in turns can explain the observed scaling relation between black holes and galaxies (Silk & Rees 1998; King 2003; Wyithe & Loeb 2003). Such a behaviour has been successfully implemented in modern cosmological simulations and the impact of active galactic nuclei (AGN) feedback on the galaxy mass content, the circumgalactic gas properties and the halo mass distribution can be dramatic (Di Matteo, Springel & Hernquist 2005; Bower et al. 2006; Croton et al. 2006; Sijacki et al. 2007; Booth & Schaye 2009; Dubois et al. 2010, 2012; Teyssier et al. 2011).

AGN feedback is supposed to have a strong impact on the stellar and gaseous content of the most massive galaxies, helping to solve the so-called 'over-cooling' problem, i.e. the expected fast cooling times and consequently high star formation rates predicted for galaxies, at odd with observational results. This was for instance highlighted by comparing theoretical models of galaxy evolution to the observed galaxy mass function (Bower et al. 2006; Croton et al. 2006). Theoretical models that do not include any form of negative feedback from stars or AGN overpredict the mass function at both the low- and high-mass end. Including stellar feedback, namely through supernova (SN) explosions, mitigates the problem at the low-mass end, but the SN energy input is not sufficient to eject gas from galaxies with virial halo velocity above a critical value of $\sim 100\,\mathrm{km\,s^{-1}}$ (Dekel & Silk 1986), nor affect the cold stream inflow in high redshift galaxies (Powell, Slyz & Devriendt 2011). The energetics involved in growing BHs are in principle comparable to,

★ E-mail: dubois@iap.fr

if not higher than, the binding energy of the host galaxy (e.g. Silk & Rees 1998; Fabian 2012). If a fraction of the rest-mass accreted energy can be tapped, it can profoundly affect the gas distribution in and around the galaxy. AGN-driven outflows can push the gas in the outskirts of the halo, even impacting the cold streams (Dubois et al. 2013; Costa, Sijacki & Haehnelt 2014). It will change the gas properties by increasing its temperature and decreasing its density, thereby, delaying the gas inflows on to galaxies and will impact the star formation rate of the host galaxy.

The masses of BHs at the centre of galaxies are observed to scale linearly with their host bulge with a BH-to-bulge mass ratio of $M_{\rm BH}/M_{\rm b} \sim 10^{-3}$. Assuming that the amount of gas mass impacted by the central AGN wind scales linearly with the bulge mass, the wind velocity $\dot{u}_{\rm w} \propto L_{\rm AGN}/M_{\rm gas} \propto M_{\rm BH}/M_{\rm b}$ is constant for BHs accreting at Eddington ($L_{\rm AGN} \propto M_{\rm BH}$). Therefore, the AGN wind velocity should be independent of the galaxy mass and AGN feedback have a non-negligible impact on high-mass galaxies as well as low-mass galaxies provided that the accretion of gas on the BH is unarrested. Measurements of the BH mass correlation with the pseudo-bulge mass show a larger scatter and fall below the correlation between BH mass and classical bulge relation (Hu 2008; Kormendy, Bender & Cornell 2011; Sani et al. 2011). A similar deficit in BH mass is also observed for low-mass BHs (Greene et al. 2010; Jiang, Greene & Ho 2011; Graham & Scott 2015). SN feedback that is an important driver of gas-mass loss in dwarf galaxies (e.g. Springel & Hernquist 2003; Dubois & Teyssier 2008; Dalla Vecchia & Schaye 2008, 2012; Hopkins et al. 2014) also turns dark matter (DM) cusps into cores (Gnedin & Zhao 2002; Mashchenko et al. 2006; Pontzen & Governato 2012; Teyssier et al. 2013), and reduces the depth of the gravitational well. Therefore, the SN feedback activity in low-mass galaxies can probably inhibit the activity of the central BH by depleting the central gas reservoir, and limiting the impact of AGN feedback. Booth & Schaye (2013) performed a numerical study of the interplay between SN and AGN feedback in a large-volume cosmological simulation (see also Sijacki et al. 2007; Crain et al. 2015; and Newton & Kay 2013 for detailed isolated and merger simulations). They found that SN and AGN feedback weaken each other's strength when jointly included in their models, compared to including only one of them at a time. However, they do not discuss whether SN feedback is able to impact the growth of BHs.

This paper is part of a series of three papers investigating the connection between SN feedback and BH growth (Paper I, this paper), SN-driven gas turbulence and BH spin evolution (Dubois et al. 2014a, hereafter Paper II), and the impact of cosmic galaxy evolution on the spin of central supermassive BHs (Dubois, Volonteri & Silk 2014b, hereafter Paper III). The main purpose of this paper is to investigate the mass evolution of the central BH of a zoomed cosmological halo at high redshift, called the Seth[1] suite of simulations, that have enough resolution to start resolving the substructures of the interstellar medium (ISM) over several Gyr. We test two implementations of SN feedback with the kinetic modelling of Dubois & Teyssier (2008), and the prescription from Teyssier et al. (2013) in order to probe how much the strength of SN activity can alter the growth of the central BH.

In Section 2 we introduce the initial conditions and the numerical models for the physics of galaxy formation, the BH mass growth and its associated AGN feedback. In Section 3 we detail our results on BH mass evolution in the Seth simulations and how SN feedback influences BH growth. Finally, in Section 4 we summarize our results.

[1] This name refers to the Egyptian god of storms, disorder and violence.

2 NUMERICAL SET-UP

The simulations are run with the adaptive mesh refinement code RAMSES (Teyssier 2002). The evolution of the gas is followed using a second-order unsplit Godunov scheme for the Euler equations. The Harten–Lax–Van Leer contact (HLLC; Toro, Spruce & Speares 1994) Riemann solver with a MinMod total variation diminishing scheme to reconstruct the interpolated variables from their cell-centred values is used to compute fluxes at cell interfaces. Collisionless particles (DM, star and BH particles) are evolved using a particle-mesh solver with a cloud-in-cell interpolation.

2.1 The Seth suite of simulations

We assume a Λcold dark matter (ΛCDM) cosmology with total matter density $\Omega_{\rm m} = 0.3$, baryon density $\Omega_{\rm b} = 0.045$, dark energy density $\Omega_\Lambda = 0.7$, amplitude of the matter power spectrum $\sigma_8 = 0.8285$, $n_{\rm s} = 0.9635$ spectral index and Hubble constant $H_0 = 68.14\,{\rm km\,s^{-1}\,Mpc^{-1}}$ consistent with the *Planck* data (Planck Collaboration XVI et al. 2014) for the initial conditions produced with MUSIC (Hahn & Abel 2011). The box size of our simulation is $L_{\rm box} = 50\,h^{-1}$ Mpc, with a coarse grid of 256^3 DM particles corresponding to a DM mass resolution of $M_{\rm res,coarse} = 8 \times 10^8\,{\rm M}_\odot$. A high-resolution region is defined around a halo of $M_{\rm vir} = 10^{12}\,{\rm M}_\odot$ at $z = 2$ that contains only high-resolution DM particles within 2 $r_{\rm vir}$ ($r_{\rm vir} = 100$ kpc) with mass $M_{\rm res,high} = 2 \times 10^5\,{\rm M}_\odot$. The halo is a progenitor of a group of galaxies whose mass is $M_{\rm vir} = 5 \times 10^{12}\,{\rm M}_\odot$ at $z = 0$.

The mesh is refined up to $\Delta x = 8.7$ pc (maximum level equals $l_{\rm max} = 21$ at redshift $z = 3$) or $\Delta x = 34.8$ pc ($l_{\rm max} = 19$ at redshift $z = 3$) using a quasi-Lagrangian strategy: when more than eight DM particles lie in a cell, or if the baryon density is larger than eight times the initial DM resolution. The minimum cell size is kept roughly constant in proper size with redshift, i.e. an additional level of refinement is added every $a_{\rm exp} = n \times 0.1$ (where $n = 1, 2, 4, 8$ and $a_{\rm exp}$ is the expansion factor of the universe) up to level 21 (or 19) at $a_{\rm exp} = 0.33$ ($z = 2$).

2.2 Physics of galaxy formation

Gas is allowed to cool by H and He cooling with a contribution from metals using a Sutherland & Dopita (1993) model for temperatures above $T_0 = 10^3$ K, which is the minimum temperature of the gas allowed through radiative losses. Heating from a uniform UV background takes place below redshift $z_{\rm reion} = 10$ following Haardt & Madau (1996). Metallicity is modelled as a passive variable for the gas advected with the flow (whose composition is assumed to be solar) and is altered by the injection of gas ejecta during SN explosions and stellar mass losses. We assume a zero initial metallicity. The gas follows an equation of state for ideal monoatomic gas with adiabatic index $\gamma = 5/3$.

The star formation process is modelled with a Schmidt law: $\dot{\rho}_* = \epsilon_* \rho / t_{\rm ff}$, where $\dot{\rho}_*$ is the star formation rate density, ϵ_* the constant star formation efficiency, and $t_{\rm ff}$ the local free-fall time of the gas. We choose a low star formation efficiency $\epsilon_* = 0.02$ consistent with observations of giant molecular clouds (Krumholz & Tan 2007) and surface density relations of galaxies (Kennicutt 1998). Star formation is allowed in regions where the gas density exceeds $n_0 = 250\,{\rm H\,cm^{-3}}$ for $\Delta x = 8.7$ pc ($n_0 = 15\,{\rm H\,cm^{-3}}$ for $\Delta x = 34.8$ pc)

and where the temperature is not larger than T_0 (corrected for the contribution of the polytropic equation of state; see below). The stellar mass resolution is $m_{\rm s,res} = n_0 \Delta x^3 \simeq 5 \times 10^3\,{\rm M}_\odot$ for the $\Delta x = 8.7\,{\rm pc}$ resolution runs ($m_{\rm s,res} \simeq 2 \times 10^4\,{\rm M}_\odot$ for the $\Delta x = 34.8\,{\rm pc}$). The gas pressure is artificially enhanced above $\rho > \rho_0$ assuming a polytropic equation of state $T = T_0(\rho/\rho_0)^{\kappa-1}$ with polytropic index $\kappa = 2$ to avoid excessive gas fragmentation. The value of n_0 for both the density threshold of star formation and the polytropic equation of state is chosen so that the Jeans length is always resolved with four cells.

Feedback from stars is taken into account assuming a Salpeter initial mass function with $\eta_{\rm SN} = 0.1$ of the mass fraction of stars ending up into a Type II supernova and releasing $e_{\rm SN} = 10^{50}\,{\rm erg\,M_\odot^{-1}}$. The numerical implementation is done in two flavours. *(i) Kinetic*: we use the implementation from Dubois & Teyssier (2008) based on a Sedov-blast wave solutions: mass, momentum and energy are deposited in the surrounding gas cells so that it mimics the Sedov solution with a mass loading of $f_w = m_{\rm ej}/m_{\rm s,\,res} = 1$, where $m_{\rm ej}$ is the gas mass of the ejecta taken from the central cell and redistributed to all the cells within the bubble of size $r_{\rm SN} = 1.5\Delta x$ (see Dubois & Teyssier 2008 for more details). *(i) Delayed cooling*: an alternative modelling is that of Teyssier et al. (2013) which locally delays the gas cooling after an explosion (in the spirit of Stinson et al. 2006), where energy is released both in the classical energy component and in a tracer component, which is denoted as 'non-thermal' and is passively advected with the flow. The passive non-thermal energy decays on a typical time-scale of $t_{\rm diss} = 0.8\,{\rm Myr}$ for $\Delta x = 8.7\,{\rm pc}$ ($t_{\rm diss} = 2\,{\rm Myr}$ for $\Delta x = 34.8\,{\rm pc}$), and the gas cooling is not allowed until the velocity dispersion associated with the non-thermal energy component is larger than $\sigma_{\rm NT} = 50\,{\rm km\,s^{-1}}$. These values of $t_{\rm diss}$ and $\sigma_{\rm NT}$ are different from the initial choice adopted in Teyssier et al. (2013) who set them up to the typical molecular cloud lifetimes (Williams & McKee 1997). Here, we prefer to adopt a resolution-dependent approach, so that the explosion is able to propagate over a few cell resolution elements. The choice of $\sigma_{\rm NT}$ and $t_{\rm diss}$ is discussed in Appendix A.

2.3 Model for BH growth and AGN feedback

We use the AGN feedback model described in Dubois et al. (2012). BHs are created at loci where both the gas density and stellar density are larger than the density threshold for star formation ρ_0 with an initial seed mass of $M_{\rm seed} = \min(0.75\rho_{\rm gas}\Delta x^3, 10^5\,{\rm M}_\odot)$ (i.e. between $[1.5 \times 10^4, 10^5]\,{\rm M}_\odot$ for $\Delta x = 34.8\,{\rm pc}$ and $[3.7 \times 10^3, 10^5]\,{\rm M}_\odot$ for $\Delta x = 8.7\,{\rm pc}$). Thus, the initial seed mass can be smaller than $10^5\,{\rm M}_\odot$ (BH mass through direct collapse; Begelman, Volonteri & Rees 2006; Spaans & Silk 2006) when gas has not accumulated enough to contain this amount of gas mass but has formed enough stars to reach the stellar density criterion. The accretion rate on to BHs follows the Bondi–Hoyle–Lyttleton (Bondi 1952) rate $\dot{M}_{\rm BH} = 4\pi\alpha G^2 M_{\rm BH}^2 \bar{\rho}/(\bar{c}_{\rm s}^2 + \bar{u}^2)^{3/2}$, where G is the gravitational constant, $M_{\rm BH}$ is the BH mass, $\bar{\rho}$ is the average gas density, $\bar{c}_{\rm s}$ is the average sound speed, $\bar{u}$ is the average gas velocity relative to the BH velocity, and α is a dimensionless boost factor with $\alpha = (\rho/\rho_0)^2$ when $\rho > \rho_0$ and $\alpha = 1$ otherwise (Booth & Schaye 2009) in order to account for our inability to capture the colder and higher density regions of the ISM. Averaged quantities are summed over four cells in radius and kernel-weighted with a Bondi radius-dependancy (see Dubois et al. 2012 for the exact form of the averaging procedure). The effective accretion rate on to BHs is capped at the Eddington accretion rate: $\dot{M}_{\rm Edd} = 4\pi G M_{\rm BH} m_{\rm p}/(\epsilon_{\rm r}\sigma_{\rm T} c)$, where $\sigma_{\rm T}$ is the Thompson cross-section, c is the speed of light, $m_{\rm p}$ is the proton mass, and $\epsilon_{\rm r}$ is the radiative efficiency, assumed to be proportional to the spin of the BH with $\epsilon_{\rm r} = 1 - E_{\rm isco} = 1 - \sqrt{1 - 2/(3r_{\rm isco})}$. Spin magnitude and direction are followed self-consistently, on the fly, in this simulation. We refer the reader to Paper II (see also Paper III for details and we just summarize here the main assumptions). The spin magnitude of BHs is modified by accretion of gas through the expression derived by Bardeen (1970). If the angular momentum of the accreted gas is misaligned with the direction of the BH spin, Lense–Thirring effect causes the innermost parts of the disc to rotate within the equatorial plane of the BH and a warped disc is created that eventually completely aligns or anti-aligns with the BH spin. The case for anti-alignment of the BH with the disc requires that (King et al. 2005) $\cos\theta < -0.5 J_{\rm d}/J_{\rm BH}$, where $\boldsymbol{J}_{\rm d}$ and $\boldsymbol{J}_{\rm BH}$ are the angular momenta of disc and BH, respectively, and θ is the angle between $\boldsymbol{J}_{\rm BH}$ and $\boldsymbol{J}_{\rm d}$. To model the unresolved accretion disc we adopt the Shakura & Sunyaev (1973) thin accretion disc solution, and define $J_{\rm d}$ to be the angular momentum calculated at smaller between-the-warp radius, i.e. the radius within which gas can effectively transfer its angular momentum to the BH within a viscous time-scale, or the self-gravity radius, i.e. the radius beyond which the disc becomes prone to fragmentation. Finally, we update BH spins after BH–BH coalescences using the analytical fit of Rezzolla et al. (2008) derived from relativistic numerical simulations of BH binaries.

In order to avoid spurious oscillations of the BH in the gravitational potential well due to external perturbations and finite resolution effects, we introduce a drag force that mimics the dynamical friction exerted by the gas on to a massive particle. This dynamical friction is proportional to $F_{\rm DF} = f_{\rm gas} 4\pi\alpha\rho(GM_{\rm BH}/\bar{c}_{\rm s})^2$, where $f_{\rm gas}$ is a fudge factor whose value is between 0 and 2 and is a function of the Mach number $\mathcal{M} = \bar{u}/\bar{c}_{\rm s} < 1$ (Ostriker 1999; Chapon, Mayer & Teyssier 2013), and where we introduce the boost factor α for the same reasons than stated above. A complementary approach, which we do not model here since gas dynamical friction is sufficient, is to explicitly model the unresolved dynamical friction exerted by DM particles on to BHs (Tremmel et al. 2015).

The AGN feedback is a combination of two different modes, the so-called *radio* mode operating when $\chi = \dot{M}_{\rm BH}/\dot{M}_{\rm Edd} < 0.01$ and the *quasar* mode active otherwise. The quasar mode corresponds to an isotropic injection of thermal energy into the gas within a sphere of radius Δx, at an energy deposition rate: $\dot{E}_{\rm AGN} = \epsilon_{\rm f}\epsilon_{\rm r}\dot{M}_{\rm BH}c^2$, where $\epsilon_{\rm f} = 0.15$ for the quasar mode is a free parameter chosen to reproduce the $M_{\rm BH}$–$M_{\rm b}$, $M_{\rm BH}$–$\sigma_{\rm b}$, and BH density in our local Universe (see Dubois et al. 2012). At low accretion rates on the other hand, the radio mode deposits the AGN feedback energy into a bipolar outflow with a jet velocity of $10^4\,{\rm km\,s^{-1}}$ into a cylinder with a cross-section of radius Δx and height $2\,\Delta x$ following Omma et al. (2004) (more details about the jet implementation are given in Dubois et al. 2010). The efficiency of the radio mode is larger with $\epsilon_{\rm f} = 1$. Though this quasar/radio mode matches the observed radiatively efficient/inefficient activities of AGN feedback (Churazov et al. 2005; Russell et al. 2013), it is possible that some accretion rate-driven jets become indistinguishable from large opening angle AGN-driven nuclear outflows after they have propagated into the multiphase gas of the galaxy (Wagner, Bicknell & Umemura 2012).

The list of the Seth simulation runs is listed in Table 1. All simulations with $\Delta x = 34.8\,{\rm pc}$ resolution are run up to $z = 2$, while the simulations with $\Delta x = 8.7\,{\rm pc}$ resolution are stopped at more intermediate redshift after the BH of the central simulation galaxy has reached self-regulation. We perform a simulation with no feedback from SNe nor from AGN (and without BH formation) denoted as SL ($\Delta x = 34.8\,{\rm pc}$). We have two simulations with SN

Table 1. List of the Seth simulation runs. First column indicates the simulation name. Second column indicates the minimum cell size. Third column indicates which SN feedback model is used. Fourth column indicates whether AGN feedback is included. Fifth column indicates the final redshift of the simulation. See the text for details.

Name	Δx (pc)	SN	AGN	$z_{\rm end}$
SL	34.8	No	No	2
SLD	34.8	Delayed cooling	No	2
SLDA	34.8	Delayed cooling	Yes	2
SHDA	8.7	Delayed cooling	Yes	2.4
SLK	34.8	Kinetic	No	2
SLKA	34.8	Kinetic	Yes	2
SHKA	8.7	Kinetic	Yes	3.5

feedback, one version with kinetic implementation SLK and another one with delayed cooling SLD, which do not include BHs ($\Delta x = 34.8$ pc). Finally, we run four simulations including AGN feedback, two with kinetic SN feedback SLKA and SHKA (respectively with $\Delta x = 34.8$ pc and $\Delta x = 8.7$ pc), and two with the delayed cooling SN feedback SLDA and SHDA.

Fig. 1 shows three projections of the gas density for the SHDA run at $z = 3.25$, one at large scale where cosmic filaments are clearly visible (within a large-scale cosmic wall), one at the halo scale where these cosmic filaments penetrate the central halo in the form of dense and cold gas mixing with the diffuse ambient medium, and another projection at galactic scale where we see the detailed structure of the gas within the galaxy in the form of dense star-forming clumps embedded into gas spiral arms. In those projections, we have indicated the virial radius $r_{\rm vir}$ of the DM halo (as the radius at which the average density in the halo is 200 times the average cosmic density). With zoom simulations, we can resolve at the same time the complex pattern of the large-scale inflows and outflows of gas into the halo and on to the galaxy, and the multiphase structure of the interstellar gas.

3 RESULTS

3.1 Quenching early BH growth with efficient SN feedback

The top panel of Fig. 2 shows the BH mass evolution for the four simulations including AGN feedback, and the bottom panel of Fig. 2 shows the BH mass accretion rate for SLDA and SLKA only. Both simulations with kinetic SN feedback, SLKA and SHKA show a rapid initial growth of the BH mass at the Eddington rate between $6 < z < 9$, with a plateau below $z < 6$ when the BH has reached self-regulation. Vice versa, both simulations with delayed cooling SN feedback, SLDA and SHDA have an extended period of time, down to $z \simeq 3.5$, where the BH grows quiescently, far from the Eddington limit. It is only below this redshift that the BH growth takes off at approximately half the Eddington rate from $2.5 < z < 3.5$ and then settles in a self-regulated state where the BH mass is comparable in runs SLDA, SHDA and SLKA. Note that the evolution of BH mass is relatively independent of resolution for the same set of SN sub-grid modelling: Eddington-limited phases appear at the same epoch and their time extensions are similar, and BH masses at self-regulation are the same to within less than a factor of 2.

At $z \simeq 3.5$, a significant galaxy merger with stellar mass ratio of $\sim$1:3 drives a large inflow of gas towards the BH. The three simulations that have reached this redshift (SLKA, SHDA and SLDA) show a rapid increase of the BH mass. In particular, the simulations with delayed cooling SN feedback have an extended near-Eddington growth for the central BH ($2.5 < z < 3.5$), due to the absence of BH self-regulation prior to the merger. Thus, the BH grows until it reaches a mass sufficiently large to pump enough energy into the gas to unbind it from the galaxy gravitational potential well.

As explained in detail in Paper II, delayed cooling SN feedback causes cold gas clumps in the galaxy to be regularly destroyed by SN explosions, every $\sim$10 Myr. Periods of intense star formation are followed by phases where stellar feedback drives large-scale outflows and hot bubbles, and removes cold gas from the galaxy at the same time. During these phases, the BH accretion rate is well below the Eddington limit,[2] while it is at Eddington when the gas cools down again and re-accumulates around the BH. This behaviour is highlighted in the inset in Fig. 2. The net effect is to overall reduce the Eddington-limited growth phase by a factor of 2, as reflected by the time-smoothed curve (for the sake of clarity) shown in the bottom panel of Fig. 2: the average BH mass accretion rate is a factor of 2 below the Eddington accretion rate between $2.5 < z < 3.5$ for the SLDA simulation (as well as for the SHDA, not represented here).

In summary, the delayed cooling implementation of SN feedback quenches the initial growth of the BH, while the kinetic SN feedback allows for a rapid Eddington-limited growth at early times. The fundamental reason is that the two different approaches to SN feedback affect the density and temperature of gas in the galaxy differently. Kinetic SN feedback imparts only momentum to the gas (with lower magnitude than the expected analytical solution; see Kimm et al. 2015). The gas quickly thermalizes if the conditions to create a shock are met, and such conditions are effectively met each time there is an SN explosion. This can be understood as follows. In the early phase, SN explosions are similar to Sedov solutions, therefore in a few time-steps, the solution will approach that of the adiabatic expansion phase, where 20 per cent of the energy is in kinetic form and 80 per cent in internal energy (e.g. Cox 1972). In dense star-forming regions, the temperature reached after an SN explosion is not high enough to avoid over-cooling of the internal component since the Sedov scale remains unresolved at our spatial resolution, and 80 per cent of the total energy of the explosion is lost, limiting the total amount of momentum that would have been transferred by the hot pressurized bubble by the end of the energy-conserving phase. Thus, if gas is not artificially prevented from cooling, dense gas can survive SN explosions, contrary to expectations. Delayed cooling SN feedback obviates this shortcoming by explicitly forcing the gas to not cool in the dense regions of star formation, mimicking the effect of internal energy deposition (however, it does not guarantee that the correct amount of momentum is imparted to the gas, only that energy is efficiently transferred to the gas.). This is the reason why simulations with delayed cooling SN feedback show that dense star-forming gas is destroyed after the ignition of the first stars into SNe (see Paper II).

Fig. 3 clearly shows these different behaviours. The gas mass within 1 per cent of the virial radius ($\simeq$100 pc at $z = 6$) has a relatively smooth time-evolution with the kinetic mode of SN feedback (SHKA, SLKA and SLK runs), comparable to the simulation

[2] Note how the Eddington accretion rate exhibits negative variations with time, while the BH mass can only increase. This apparent paradox is explained by variations in the spin parameter of the BH which in turn change the Eddington accretion rate by modifying the radiative efficiency $\epsilon_{\rm r}$ (see Papers II and III for a complete overview of the BH spin evolution).

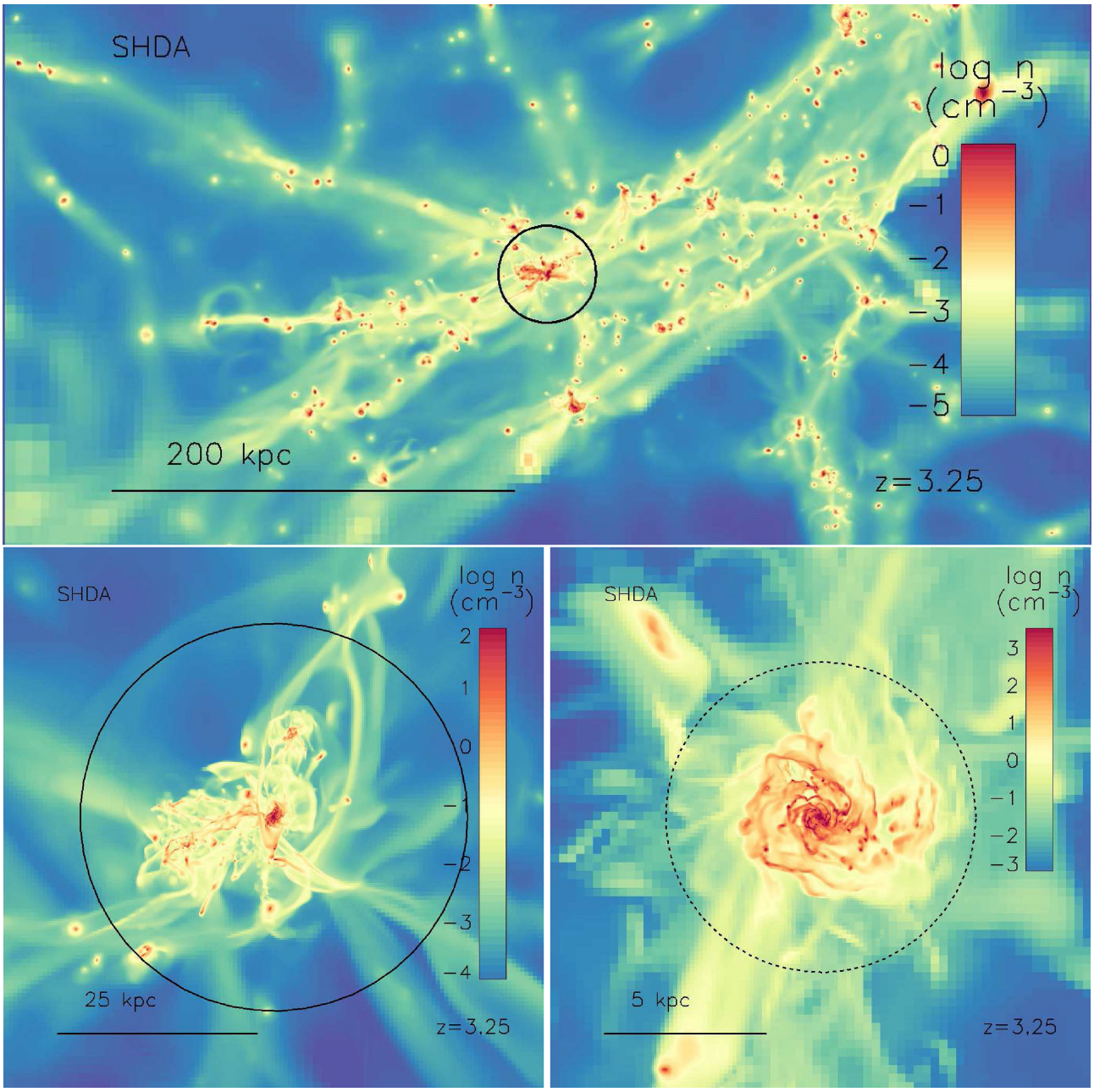

Figure 1. Projected gas densities of the Seth simulation (SHDA run) for the zoom region (top panel), the central halo (left panel), and the central galaxy (right panel) at $z = 3.25$. The solid circle indicates the virial radius $r_{\rm vir}$ of the DM halo, and the dashed circle is for $0.1r_{\rm vir}$. We can see several cosmic cold filaments that connect to the central halo with a complex gas distribution once they reach the central galaxy. The galaxy clearly exhibits a rotating disc-like structure with dense star-forming clumps of gas.

without SNe (SL run). Simulations with delayed cooling SN feedback (SHDA, SLDA and SLD runs) instead show strong variations in the amount of gas mass relatively close to the BH. These variations are an intrinsic feature related to the modelling of SNe as they are detected independently of the presence of the AGN feedback and the choice of resolution.

Fig. 4 highlights two specific times representative of the two phases characteristic of the delayed cooling SN feedback: intense star formation driven by cold gas followed by SN activity which in turn removes the cold dense gas. The four left-most panels of Fig. 4 show the projected gas density at a maximum ($z = 5.80$) and near a minimum ($z = 5.65$) of the gas mass within 1 per cent of the halo virial radius for the SHDA simulation (cf. Fig. 3, top panel). The dense gas in the galaxy appear edge-on with some cold clumps near the very centre of the galaxy at $z = 5.8$. However, at $z = 5.65$, the dense gas has disappeared from the centre of the image due to SN explosions that have disrupted this cold gas component. For comparison, the four right-most panels of Fig. 4 show the same gas density projections for the SHKA simulation, where a thin disc of gas also appears nearly edge-on, and this disc is present in both time-steps. Therefore, contrary to the case with delayed cooling SN feedback (SHDA), the simulation with kinetic SN feedback (SHKA) does not show any sign of the destruction of the cold star-forming gas in the galaxy. Finally, the presence at all times of this

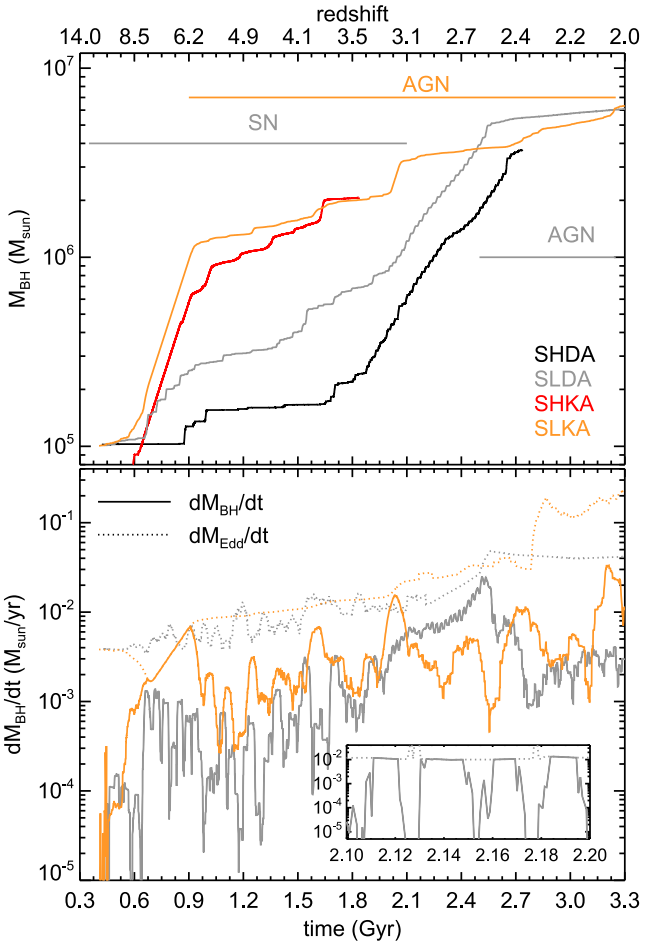

Figure 2. Top: BH mass as a function of time (redshift) for SHDA (black), SLDA (grey), SHKA (red) and SLKA (orange). Bottom: BH mass accretion rate (solid lines) and Eddington mass accretion rate (dotted lines) for SLDA and SLKA simulations only and values are smoothed for the sake of readability with a 5 Myr time window. The small inset shows the exact BH mass accretion of the SLDA simulation between $2.1 < t < 2.2$ Gyr. The horizontal lines indicate the SN and AGN regulation phases of the BH growth with a colour that corresponds to the simulation it refers to. The simulations with kinetic SN feedback (SLKA and SHKA) show a rapid growth of the BH mass at early times and quickly regulate the BH mass growth. The simulations with delayed cooling SN feedback (SLDA and SHDA) instead inhibit the early BH growth and delay its Eddington-limited phase to later times (below redshift $z < 3.5$).

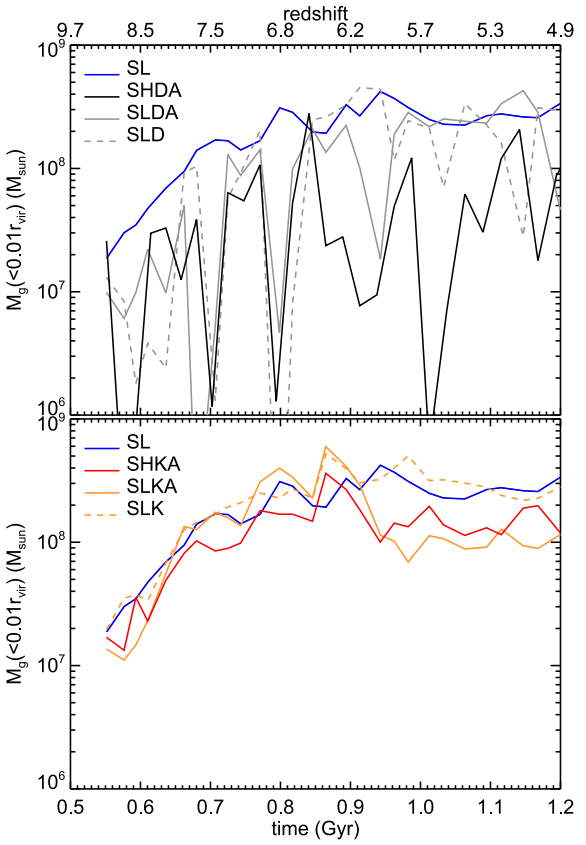

Figure 3. Mass of gas within 1 per cent of the halo virial radius as a function of redshift for SHDA (black), SLDA (grey solid), SLD (grey dashed), SHKA (red), SLKA (orange solid), SLK (orange dashed), and SL (blue). Only simulations with delayed cooling SN feedback exhibit rapid variations in the gas mass content at early times due to the efficient evacuation of the cold gas from the galaxy centre.

cold gas reservoir near the BH for SHKA (and SLKA) is the reason for the rapid early growth of the BH, while in SHDA (and SLDA) disruption of dense gas is efficient enough to starve the BH.

Fig. 5 shows the gas number density in the vicinity of the BH as a function of time, as well as the gas sound speed and its relative velocity with respect to the BH. During the early growth of the BH in the SLKA run ($6 < z < 9$), the gas density near the BH is above $n > 10^3\,\mathrm{cm}^{-3}$, the gas sound speed is $\sim$50 km s^{-1}, and the gas relative velocity is $u < 10\,\mathrm{km\,s}^{-1}$. Therefore, the gas is sufficiently dense, cold, and with very little turbulence (relative gas motion) to allow the Bondi accretion rate to be Eddington-limited, and the BH to grow to reach self-regulation in $\sim$ 200 Myr. During the out-of-Eddington phases, gas density drops to lower values ($n \simeq$ 20–100 cm^{-3}), the gas sound speed increases ($c_s \simeq 200\,\mathrm{km\,s}^{-1}$), as does the gas relative velocity ($u \simeq$ 100–300 km s^{-1}), all being consequences of the energy input from the AGN while the BH is at self-regulation. It is particularly striking that these quantities change abruptly (gas density increases and velocities decrease) with the new inflow of gas triggered by the merger at $z \simeq 3.5$ ($t \simeq 2$ Gyr).

For the SLDA run, gas density, sound speed and gas relative velocity values up to $z = 3.5$ are comparable to the values obtained in the SLKA run during the BH self-regulation phase (i.e. after $z \sim 6$; $t \simeq 1$ Gyr). The main difference is that, in the SLDA run, the absence of Eddington-limited phase is due to the energy input from SNe that destroy the dense star-forming gas, keeping the density low, the temperature high and the ISM turbulent. The gas number density in SLDA between $5 < z < 9$ shows rapid variations with peaks at $n \simeq 500\,\mathrm{cm}^{-3}$ followed by drops to $n \simeq$ 50–100 cm^{-3}. It illustrates well the 'breathing' of the galaxy when delayed cooling SNe feedback is employed (see also Fig. 3): dense gas accumulates in the galaxy (and near the BH) for $\Delta t \sim 10$ Myr until the first SNe explode and drive the cold gas in an expanding wind that falls back on to the galaxy in a dynamical time. At $z \simeq 3.5$, a merger triggers fresh gas near the BH that accumulates and leads to the near-Eddington phase: gas sound speed and gas relative velocity are still high (respectively $\simeq$300 and 100 km s^{-1}), but the gas density remains steadily at $n \simeq 500\,\mathrm{cm}^{-3}$. Thus, even though velocities are on average large due to the starburst activity induced by the merger, the gas density is sufficiently high to keep the BH growing near the Eddington rate. Note that these quantities have strong small-scale time variations (on a time-scale of 10 Myr) that cannot be represented in Fig. 5. After the merger-driven triggering of the BH

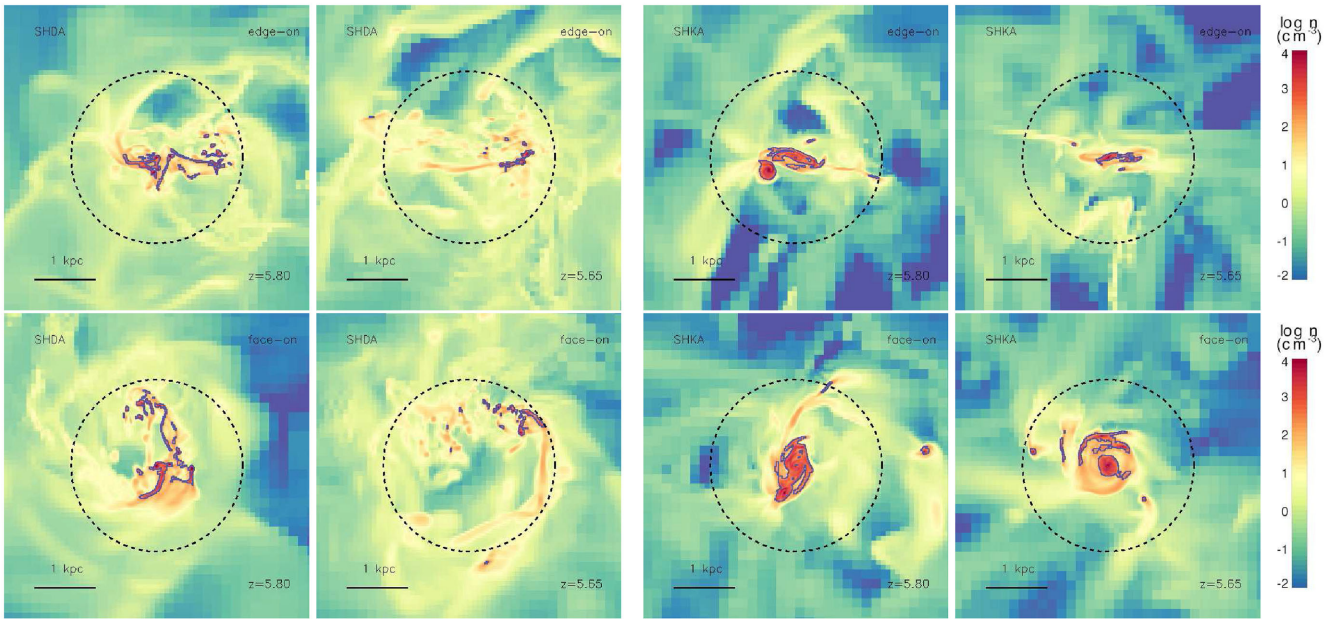

Figure 4. Gas density projections of the central galaxy at $z = 5.8$ (first and third columns) and $z = 5.65$ (second and fourth columns) for the SHDA (four left-most panels) and SHKA (four right-most panels) simulations, seen edge-on (first row) or face-on (second row). The blue contours surround the regions of star formation with gas density above 250 H cm^{-3}. The dashed circles indicate the 10 per cent virial radius of the halo. In the simulation with the delayed cooling SN feedback, dense star-forming clumps in the centre of the galaxy are regularly destroyed and reprocessed, while in the simulation with kinetic SN feedback the star-forming gas settles into a long-lived compact rotating disc.

growth in the SLDA simulation, the BH ends up self-regulating at $z \simeq 2.5$, and the gas density drops, while the gas sound speed and relative velocity increase to values comparable to that of the SLKA simulation at the same epoch and for the same reason: its AGN activity is strong enough to remove the cold gas from its direct surroundings.

3.2 Why SNe are able to quench BH growth in low-mass galaxies

From our set of simulations, we have seen that the central BH growth is initially inhibited by the presence of SNe when they are associated with a delayed cooling prescription that maximizes the gas motion. Now, we want to compare the theoretical expectation of the velocity of SN-driven winds to the escape velocity of the galactic bulge. Therefore, if SN-driven winds are efficient, they should drive the gas out of the bulge (and star-forming clumps), until the bulge becomes sufficiently massive to retain any outflows in its sphere of influence.

The characteristic velocity of SN-driven winds in a star-forming clump is of the order of the Sedov velocity:

$$u_{\rm SN} \simeq 1.2\sqrt{\frac{m_{\rm new,s} f_{\rm blast} \eta_{\rm SN} e_{\rm SN}}{m_{\rm g}}} \simeq 270\sqrt{\frac{f_{\rm blast}\eta_{\rm SN}}{0.1}}\sqrt{\frac{(m_{\rm new,s}/m_{\rm g})}{0.1}}\,{\rm km\,s^{-1}}\,, \quad (1)$$

where $e_{\rm SN} = 10^{50}$ erg M$_\odot^{-1}$ is the SN type II specific energy, $f_{\rm blast} \leq 1$ is the fraction of SN energy input that is *effectively* transferred into the gas (in the case of efficient energy losses through gas cooling $f_{\rm blast}$ can be much lower than 1), and $m_{\rm new,\,s}$ is the amount of new stars formed within the gas clump. $m_{\rm new,\,s}$ depends on the clump free-fall time (i.e. the clump gas density) and the time-scale before the ignition of the first SNe $t_{\rm SN} = 10$ Myr. Integrating the Schmidt law over time and volume, and assuming that the gas density is nearly constant over $t_{\rm SN}$, one obtains

$$\frac{m_{\rm new,s}}{m_{\rm g}} \simeq \epsilon_* \frac{t_{\rm SN}}{t_{\rm ff}}\,, \quad (2)$$

where the gas free-fall time $t_{\rm ff}$ for the central gas clump of mass $\sim 10^8$ M$_\odot$ in the proto-bulge (see Fig. 3) and of $\sim$100 pc radius[3] is $\simeq$1.6 Myr. Therefore, the ratio of $m_{\rm new,\,s}/m_{\rm g} \simeq 0.12$. To summarize, the velocity of the blast propagating into the central gas clump (bulge) should be close to 270 km s^{-1} for the delayed cooling case, where $f_{\rm blast} \sim 1$ is ensured by construction.

The escape velocity of a clump is

$$u_{\rm esc} = \sqrt{\frac{2\,G m_{\rm cl}}{r_{\rm cl}}}\,, \quad (3)$$

where $m_{\rm cl}$ and $r_{\rm cl}$ are, respectively, the clump total mass and clump radius. If we replace the clump radius $r_{\rm cl}$ by the bulge radius $r_{\rm b} \simeq 100$ pc, and the clump mass $m_{\rm cl}$ by the bulge mass $M_{\rm b} \simeq 10^9$ M$_\odot$ that are the representative values when the BH starts growing at the Eddington rate, the escape velocity is, therefore, $u_{\rm esc} \simeq 300$ km s^{-1} and is comparable to the Sedov velocity produced by SN winds in the bulge. Thus, below $M_{\rm b} \lesssim 10^9$ M$_\odot$, and provided that $f_{\rm blast} \sim 1$, the gas cannot accumulate in the bulge and is regularly ejected from the bulge gravitational potential well (see the top panel of Fig. 3 for SHDA, SLDA and SLD), and above $M_{\rm b} \gtrsim 10^9$ M$_\odot$ the gas is locked into the bulge, in a fountain that ejects gas that quickly falls back, and provides the fuel to feed the central BH over a prolonged duration.

[3] This radius is obtained doing a bulge–disc decomposition on the stellar density profile of SHDA using a fitting procedure with two exponentials. The bulge scale radius is of $\sim$25 pc. We use a radius of 100 pc in order to account for 90 per cent of the stellar mass content in the bulge in that case.

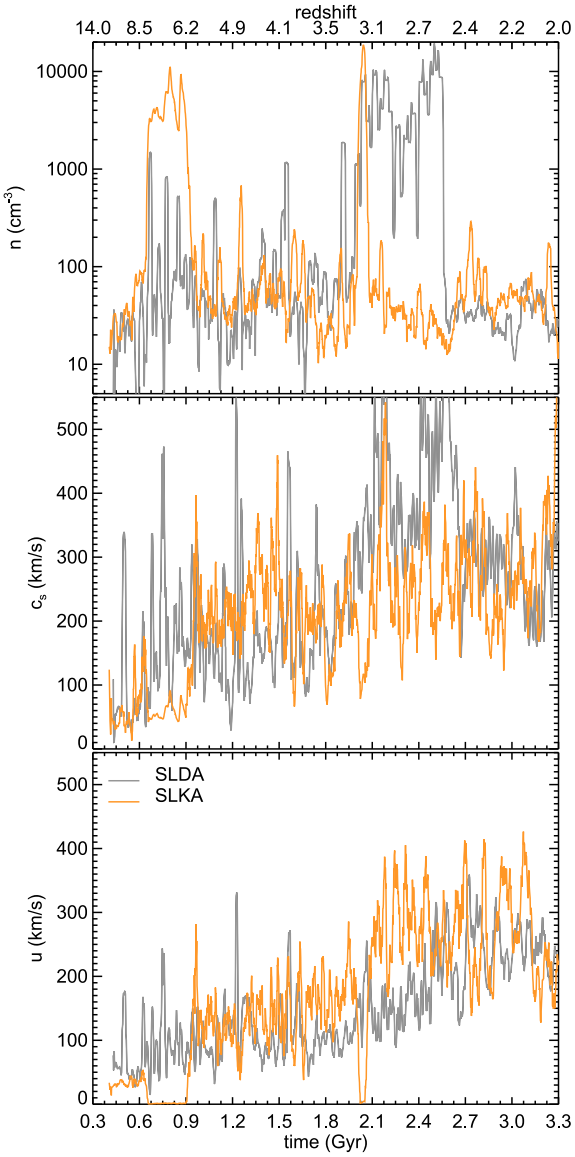

Figure 5. Gas number density (top), gas sound speed (middle), and gas relative velocity (bottom) near the BH as a function of time (redshift) for SLDA (grey) and SLKA (orange).

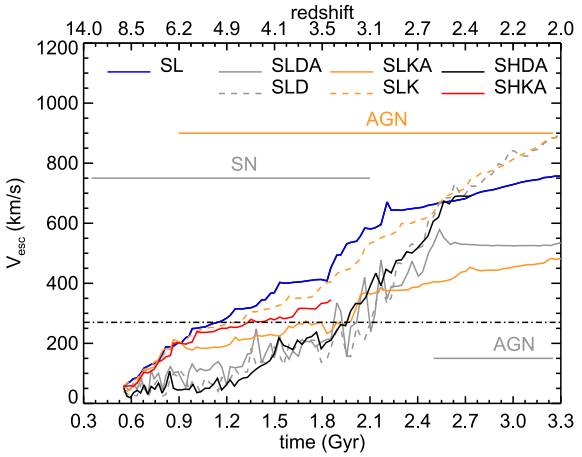

Figure 6. Escape velocity from the bulge (at $r_b = 100$ pc from the galactic centre) as a function of time (redshift) for SL (blue), SLDA (grey solid), SLD (grey dashed), SLKA (orange solid) and SLK (orange dashed). The horizontal dot–dashed line corresponds to the velocity of SN-driven winds in the bulge with $f_{blast} = 1$ (corresponding to the delayed cooling SN model). Grey horizontal lines mark the SN- and AGN-dominated evolution in the delayed cooling case. The orange horizontal line marks AGN-dominated evolution in the kinetic feedback case. In SLDA and SLD, the end of the SN-regulated growth of the BH corresponds to the time when the escape velocity of the bulge exceeds the velocity of SN-driven wind. In the SLKA and SLK runs, the BH grows as soon as it is created, and there is no SN-dominated phase, which indicates that the SN-driven winds for the kinetic model have a very low energy-conversion efficiency ($f_{blast} \ll 1$) and, thus, an ejection velocity below the horizontal dot–dashed line.

Fig. 6 shows the escape velocity within 100 pc distance from the centre of the galaxy. For the sake of simplicity, we call that region the bulge here, but note, however, that the bulge is not always well defined, especially at early times when only short-lived star-forming clumps are present in the galaxy with delayed cooling SN feedback. At early times ($z > 3.1$), since SNe in the delayed cooling cases prevent the gas from accumulating too rapidly in the bulge, we observe that the escape velocity of the bulge increases more slowly than in the kinetic feedback cases. At $z = 3.1$, the escape velocity in the bulge in SLD and SLDA reaches 270 km s^{-1}. This value is the predicted SN-driven wind velocity in the bulge for $f_{blast} = 1$ and marks the end of the SN-driven evolution of the bulge gas content. From this moment, the BH can grow close to the Eddington rate and settles into self-regulated growth ∼400 Myr later at $z = 2.6$. Note that escape velocities in the bulge in both SLD and SLDA, respectively with and without AGN feedback, are similar during the SN-driven evolution since the AGN has as yet no impact on the surrounding gas as SN-driven winds prevent the BH from accreting gas at near-Eddington. There is a difference in escape velocity between SLDA and SHDA (i.e. when changing resolution) at late times. This is attributed to the more compact bulge in the high-resolution run (∼25 pc against ∼60 pc bulge scale radius), although the mass of the bulge remains identical (see section 3.3).

For the kinetic cases, SLK and SLKA, the bulge endures a rapid mass growth already very early on due to the inefficiency of the coupling of the SN energy to the gas ($f_{blast} \ll 1$). Until the BH has reached its self-regulated phase in SLKA, the bulge mass can grow steadily, as it does in the case without any feedback from SNe nor AGN (SL run). Finally, the growth of the bulge escape velocity (or equivalently the bulge mass) is suppressed once the BH is massive enough to have a significant impact on the surrounding gas, which is around $z \simeq 6$ in SLKA. Interestingly, the bulge escape velocities of SLDA and SLKA are similarly $u_{esc} \simeq 500$ km s^{-1} once both simulations have reached the BH self-regulated phase ($z < 3$), and their value is smaller compared to the runs without AGN feedback $u_{esc} \simeq 800$ km s^{-1} (SL, SLD and SLK).

A more in-depth investigation of the structure of the gravitational potential well, and, thus, mass distribution, is shown in Fig. 7, where we see how the gas mass accumulates differently in SLKA and SLDA at early epochs. In the SLKA run, the escape velocity peaks at ∼400 km s^{-1} at $z = 4.71$ within the galactic bulge ($<0.01\, r_{vir} = 200$ pc) due to the high concentration of stars in the central region of the galaxy (and with a low concentration of gas and DM). Vice versa, in the SLDA run at the same redshift, the peak of the escape velocity is 250 km s^{-1} at a distance $\simeq 0.05\, r_{vir} = 1$ kpc that roughly corresponds to the galaxy radius. The escape velocity within the galactic bulge reaches lower values around 150 km s^{-1}, a factor of 2 below the SN wind velocity. In this simulation, the dominant mass component in the bulge is the gas with a fraction above 50 per cent. Though the amount of gas within the bulge is

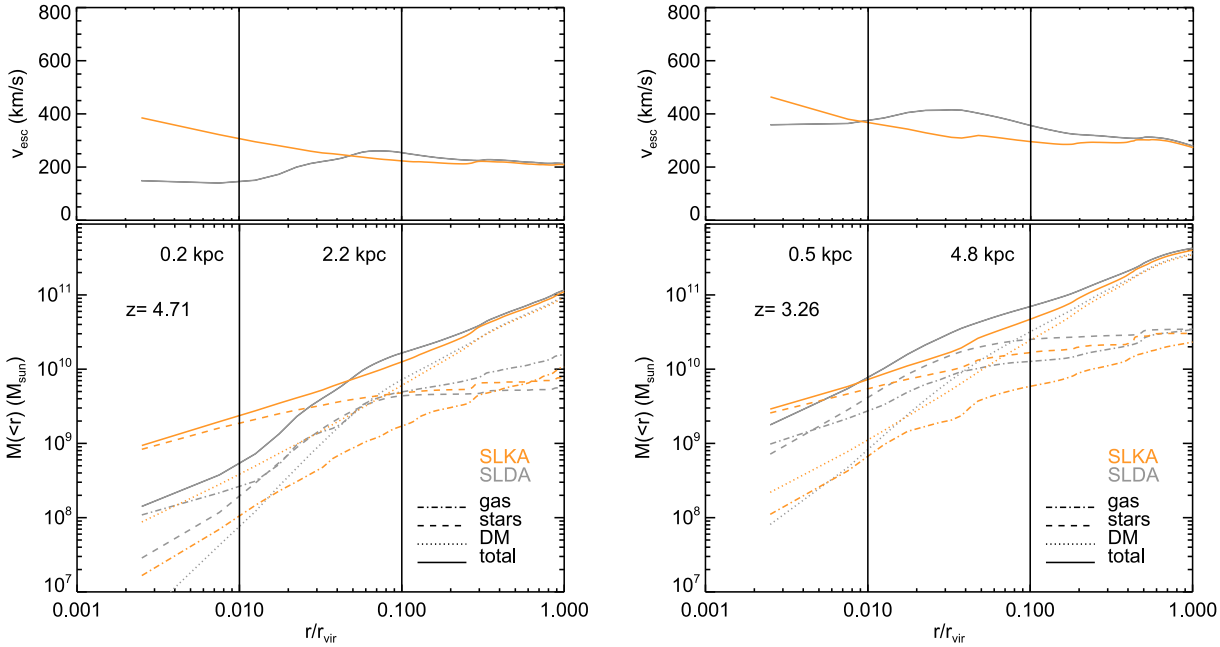

Figure 7. Integrated mass profiles as a function of radius normalized to the halo virial radius at $z = 4.71$ (left panel) and $z = 3.26$ (right panel). Grey lines are for the SLDA simulation and orange lines are for the SLKA simulation. Solid lines are the total masses, dotted lines are the DM masses, dashed lines are the stellar masses and dot–dashed lines are the gas masses. At early times $z = 4.71$, the simulation including SN feedback with delayed cooling (SLDA) has reduced the mass concentration within 10 per cent of the virial radius compared to the simulation with kinetic mode of SN feedback (SLKA).

larger by a factor of 2 in SLDA than in SLKA, the amount of total mass is lower in SLDA because of the frequent expulsions of gas, which translate into a much lower amount of stellar mass. Note also the strong enhancement of the DM concentration within the central region in the SLKA run due to adiabatic contraction (Gnedin et al. 2004).

At the near-Eddington growth of the BH in SLDA ($z = 3.26$), the escape velocity profile in the galaxy has been boosted by a factor of $\sim$2, reaching $\sim$400 km s^{-1} at $r \simeq 0.03\, r_{\rm vir}$ (350 km s^{-1} at the galaxy centre). These new values of the escape velocity in SLDA create a gravitational barrier large enough for the SN winds to be trapped within the core of the galaxy, and, therefore, allow maintenance of a steady reservoir of cold gas for the BH to grow from. The SLKA escape profile shows a small variation with the $z = 4.71$ profile; it is simply rescaled by a $\gtrsim$1 factor. Although the stellar mass and DM mass content within the galaxy bulge are now similar in SLKA and SLDA, the amount of gas is still a factor of 5 different, and the gas fraction in SLDA is close to 30 per cent while it is below 10 per cent in the SLKA.

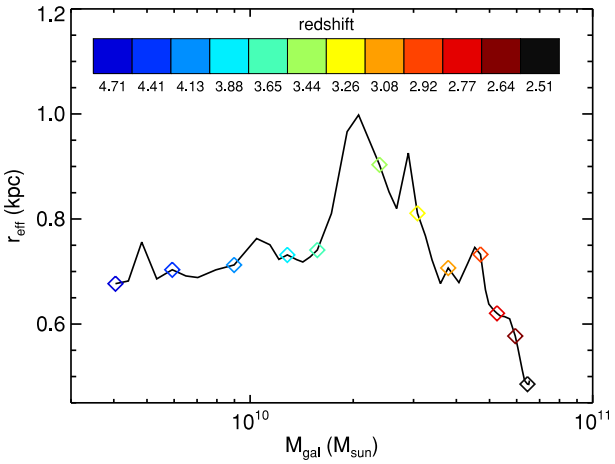

Figure 8. Evolution with redshift of the effective radius and the stellar mass of the central galaxy in SHDA at different redshifts (indicated by the colours). The 1:3 merger is happening from $z = 3.7$ to 3.5 when the effective radius quickly increases and as quickly contracts.

3.3 The triggering of nuclear inflows by a major merger

In the simulations with delayed-cooling SN feedback, the BH enters its Eddington-limited growth phase once the galaxy encounters a significant merger with a stellar mass ratio of $\sim$1:3. Fig. 8 shows the galaxy effective radius in SHDA as a function of its stellar mass for different redshifts. The stellar mass of a galaxy is defined by the total amount of stars which belong to the same structure using a galaxy finder algorithm. We employ the AdaptaHOP finder (Aubert, Pichon & Colombi 2004; Tweed et al. 2009) on to the distribution of all stars in the simulation. For computing the effective radius, we project the stellar mass along the three Cartesian axes of the box, measure the half-mass projected radius of each projection, and take the geometrical average of the three components. In the early stage of the galaxy evolution ($z > 4$) its effective radius stay nearly constant even though the galaxy stellar mass increases. Once the 1:3 merger happens from $z = 3.7$ to $z = 3.5$ the galaxy radius quickly increases, but then contracts further.

Fig. 9 illustrates further the connection between the BH feeding and the morphological transformation of the galaxy in SHDA. This figure shows (at the top) the BH growth relative to Eddington as a function of redshift with the two largest merger mass ratios highlighted in red, and the panels show the stellar emission in rest-frame *u, g, i* filters at various redshifts. We see that above $z > 3.6$ the galaxy is a very disturbed object with no bright galaxy bulge, plenty of blue stellar clusters and a disc scale radius almost constant. Between $3.21 < z < 3.6$, a stellar disc forms with a central bar that triggers

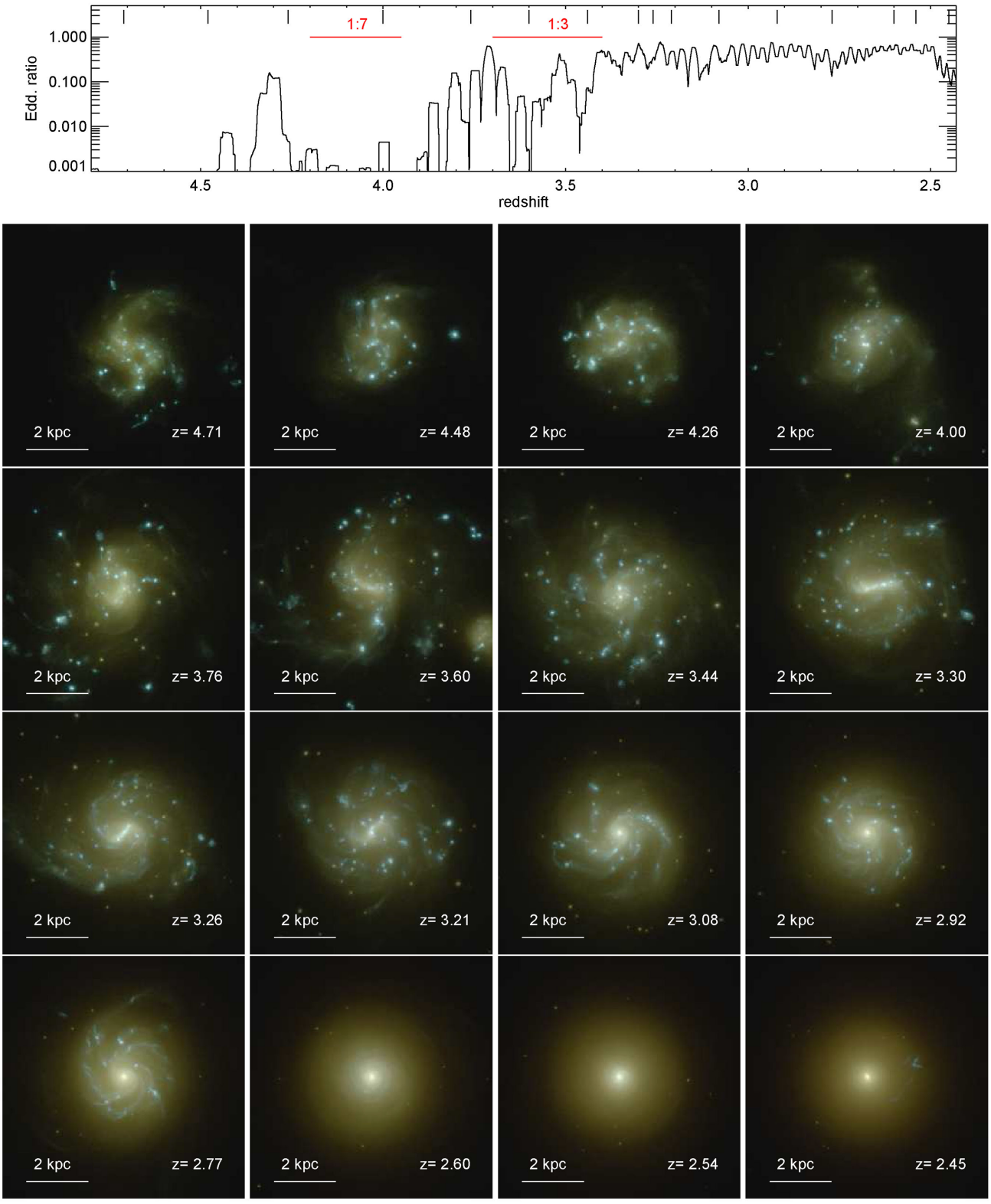

Figure 9. Top: Eddington ratio of the central BH as a function of redshift for the SHDA. Important mergers for the BH growth are indicated as the red horizontal lines with their stellar mass ratio. Small vertical bars pinpoint the redshifts of the stellar images. Bottom: stellar emission of the central galaxy at different redshifts for SHDA as they would be observed face-on through $u - g - i$ filter bands in the rest frame. Extinction by dust is not taken into account. We see three main phases of galaxy evolution: the morphologically disturbed phase ($z > 3.6$), the disc settling phase with a bar ($2.77 \leq z \leq 3.6$), and the red nugget ($z \leq 2.6$).

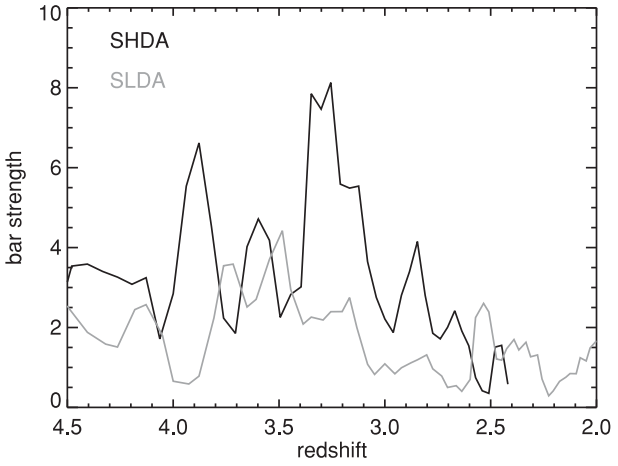

Figure 10. Evolution with redshift of the bar strength of the central galaxy in SHDA (black) and SLDA (grey). The bar becomes stronger between $3.1 < z < 3.4$ after the major merger (occurring at $3.4 < z < 3.7$) in SHDA, and is the signature of a nuclear inflow of gas that triggers the steady Eddington-limited growth of the central BH at this redshift.

the inflow of gas in the centre of the galaxy resulting in a compact bulge later on at $z < 3.21$. In this redshift range, the galaxy settles into a disc-like shape and has increased its radius (see also Fig. 8) at the onset of the merger with its massive companion (that one can see at the east/south-east of the panel corresponding to $z = 3.6$). We suggest that the central bar and the resulting nuclear inflow are induced by the new gravitational torques exerted by the merger, which explain the rise in the BH growth at the near-Eddington rate at this redshift. Finally, this merger has allowed sufficient amounts of mass to concentrate around the BH through new gas inflows. In return, the gravitational barrier (or escape velocity) of the bulge has become large enough to prevent the SN winds from depleting the cold gas content in the centre of the galaxy. The consequence for the BH is that there is now plenty of gas residing in the bulge, and the BH can, from $z = 3.6$, grow at the near-Eddington rate until it reaches self-regulation $z \simeq 2.5$. At the end of the merger ($z = 3.4$), the galaxy starts to shrink and later on ($z = 2.6$) form a red nugget with a smooth red disc component and a compact stellar bulge. In Fig. B1, we show the morphological evolution of the central galaxy in the SLDA run extended to lower redshift $z = 2$. The morphological evolution is remarkably comparable to the high-resolution run SHDA, and shows the formation of a spiral blue disc galaxy below the redshift of $z < 2.5$.

In Fig. 10, we show the bar strength of simulations SHDA and SLDA. The bar strength is obtained by measuring the inertia tensor of young stars (younger than 10 Myr) within 1 kpc radius from the centre of the galaxy. The inertia tensor is then diagonalized to get the three principal axis of the ellipsoid $e_1 < e_2 < e_3$. We, then, use those three axis to evaluate the strength of the asymmetric distribution of young stars (i.e. the so-called 'bar') with the bar strength being equal to $(e_3/e_2 - 1)e_3/e_1$. We see that the bar strength in SHDA has two main peaks. The first one at $z \simeq 3.9$, which correspond to the triggering of accretion bursts on to the central BH, though they are discontinuous and not at Eddington (see top panel of Fig. 9 between $3.6 < z < 3.9$). The second more extended peak is at $3.1 < z < 3.4$, and that is the result of the major merger occurring before at $3.4 < z < 3.7$ and which is responsible for the global destabilization of the disc and the central nuclear inflow. This nuclear inflow takes the form of a bar that can be seen in the stellar light distribution in Fig. 9. Note that the same bar strength

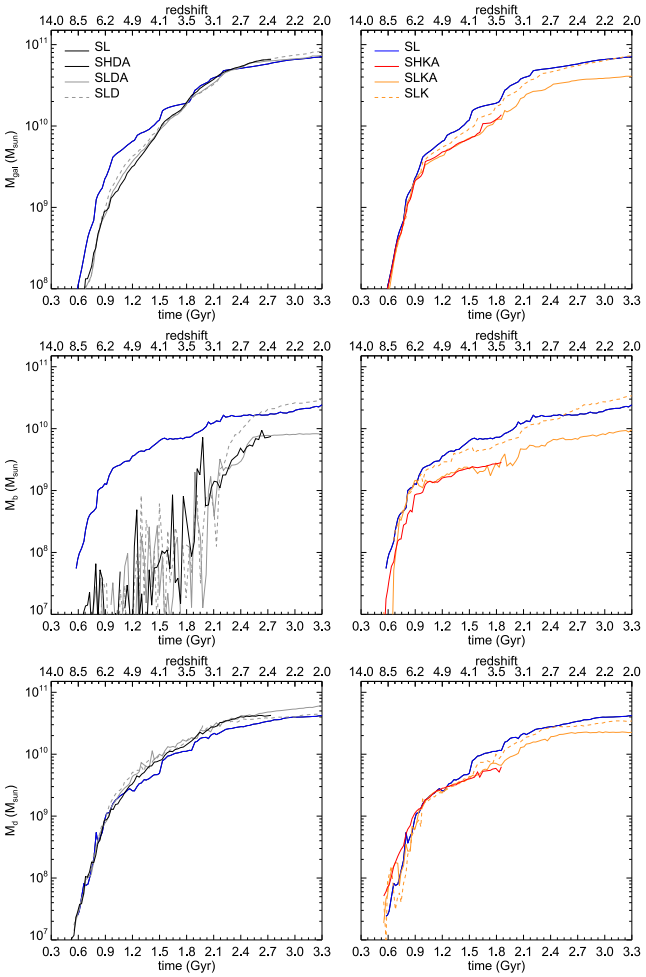

Figure 11. Galaxy mass (top panels), bulge mass (middle panels), and disc mass (bottom panels) as a function of time (or redshift). In the presence of delayed cooling SN feedback (SLD, SLDA and SHDA) the bulge is suppressed at early times (above $z > 3$). The presence of kinetic SN feedback has little impact on galaxy mass, both bulge and disc (SLK compared to SL). However, the presence of the AGN greatly reduces the bulge mass content independently of both the SN implementation and resolution.

increase is observed for the lower resolution run SLDA, except that it appears slightly earlier in time ($3.1 < z < 3.8$) with a lower amplitude (that can probably be attributed to resolution effects). It also explains why the Eddington-limited growth in SLDA appears slightly earlier than in SHDA (see top panel of Fig. 2).

Fig. 11 shows the galaxy stellar mass as a function of time for the different simulations together with the bulge and disc mass content. To compute the bulge and disc mass, we build the one-dimensional stellar surface density profile extracted from the galaxy finder (in order to remove stellar sub-structures using ADAPTAHOP), and fit two exponentials $\Sigma_{b,d} = \Sigma_{0,b,d}\exp(-r/r_{b,d})$, where $\Sigma_{0,b,d}$ and $r_{b,d}$ are the free parameters of the fit. The bulge M_b and disc M_d masses are simply $M_{b,d} = \Sigma_{b,d} 2\pi r_{b,d}^2$, where the bulge is represented by the profile with the smaller r_b. The galaxy mass is smaller in all delayed cooling SN feedback runs (SLD, SLDA, SHDA) at early times compared to the no feedback run (SL), until $z \simeq 3$, but this is essentially due to the suppression of the bulge component. The disc component in SL, SLDA and SHDA is comparable and somewhat larger than the disc component of the SL run.

The kinetic mode of SN feedback, instead, has very little impact on the stellar mass content, for both the bulge and disc: the mass evolution of the SLK and SL is very close. We see that the bulge mass in the SLDA and SHDA runs finally settles at $z = 3$ once the major merger has happened at $z = 3.5$. Before the settling of the bulge, its mass shows extreme variations with time. This effect is the consequence of the quick removal of gas due to SN explosions, creating rapid variations in the bulge gravitational potential well and leading to the flattening of the stellar mass distribution (Governato et al. 2010) with the destruction of the central stellar clump. Note that the presence of the AGN feedback has a non-negligible effect on the mass content of the bulge. Once the BH is self-regulating ($z \simeq 6$ for SLKA and SHKA, and $z \simeq 2.5$ for SLDA and SHDA), the bulge mass growth is regulated by the presence of the AGN feedback with more than a factor of 2 difference at $z = 2$ compared to the simulations without AGN.

At $z = 2$, the halo mass is $M_{\rm h} = 10^{12}\,{\rm M}_{\odot}$, and the expected stellar mass at this redshift obtained from abundance matching is $M_{\rm gal} \simeq 2 \times 10^{10}\,{\rm M}_{\odot}$ (Moster, Naab & White 2013). In all our simulations, the central galaxy ends up being more massive than expected from semi-empirical predictions by a factor of 2 ($M_{\rm gal} \simeq 4 \times 10^{10}\,{\rm M}_{\odot}$) for SLKA and by a factor of 3.5 ($M_{\rm gal} \simeq 7 \times 10^{10}\,{\rm M}_{\odot}$) for SLDA. This tension is possibly due to the missing extra sources of stellar feedback including stellar winds and radiation pressure from young stars (Hopkins, Quataert & Murray 2011). Including more stellar feedback and having a larger star formation efficiency should reduce the stellar mass content of the galaxy and delay the formation of the bulge (Governato et al. 2010; Roškar et al. 2014; Agertz & Kravtsov 2015), and, as a consequence, further inhibits the growth of the central BH. It is worth noting that the SN-quenching of BH growth in SLDA occurs until the halo reaches a mass of $M_{\rm h} \simeq 3 \times 10^{11}\,{\rm M}_{\odot}$ at $z \simeq 3$, which is smaller than the expected peak of the stellar conversion efficiency at $z = 3$ ($M_{\rm h,peak} \simeq 3 \times 10^{12}\,{\rm M}_{\odot}$; Moster et al. 2013). Therefore, if more realistic stellar feedback can produce lower stellar mass at a corresponding halo mass before that peak, we can assume that the halo mass at which this SN-quenching of BH growth mechanism becomes ineffective appears for more massive haloes, probably around the peak of stellar conversion efficiency.

In the case with delayed cooling SN feedback, the BH starts to grow (see Fig. 2) when the bulge definitively settles at $z \sim 3.5$, while in the kinetic feedback case bulge and BH start growing earlier on. In Fig. 12 we compare, in the top panel, the ratio between BH and bulge mass as a function of time. In SLKA and SHKA a roughly constant ratio, $\simeq 7 \times 10^{-4}$, is maintained from early times, $z \sim 6$ all the way to the end of the simulations. This value is only slightly lower than the ratio found at $z = 0$, 1–3 $\times\ 10^{-3}$. SHDA and SLDA present instead a very different evolution: the BH's growth is very slow until $z \sim 3.5$, and the bulge has not settled yet. Therefore, before that time, the ratio between BH and bulge mass varies rapidly, despite the BH mass remaining roughly constant. In this case a measure of the BH-to-bulge ratio would indicate a value typically higher than $\sim 10^{-3}$. The bulge is able to grow, and to let the BH grow, only after the potential of the bulge has reached a large enough value (cf. Section 3.2). We have also evaluated the ratio of the BH to total galaxy mass (bottom panel of Fig. 2). In fact, except for a few exceptions (Peng et al. 2006; Schramm & Silverman 2013) high-z measurements of the masses of quasar and AGN hosts cannot perform a robust bulge-to-disc decomposition, and often the total stellar mass is measured (but compared to the BH–bulge relationship at $z = 0$). In SLKA and SHKA the ratio gently decreases with cosmic time, highlighting that the total stellar mass grows faster than the BH. Therefore, in the SLKA and SHKA

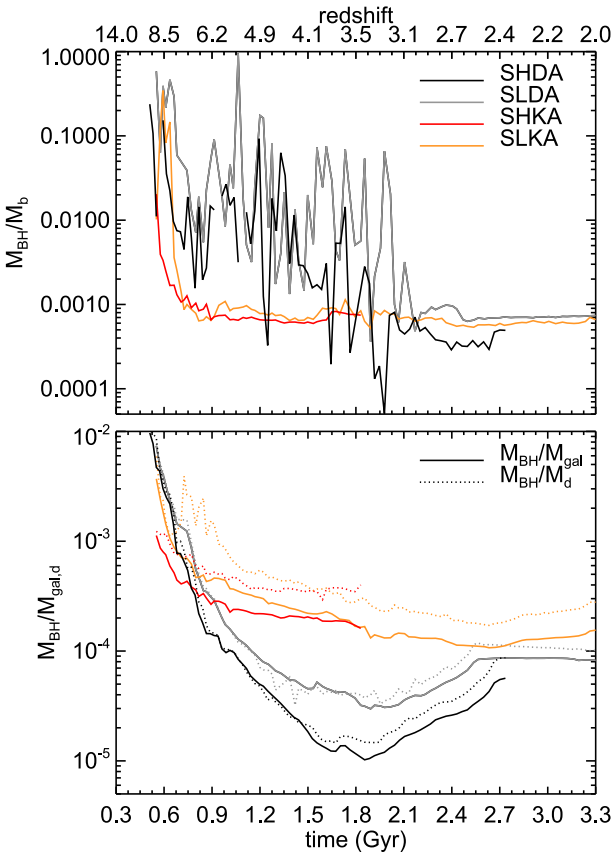

Figure 12. BH mass over bulge mass (top panel), BH mass over galaxy mass (solid) or disc mass (dotted, bottom panel) as a function of time (or redshift) for the SHDA (black), SLDA (grey), SHKA (red), and SLKA (orange) simulations. The BH over bulge mass ratio is $\simeq 7 \times 10^{-4}$ for all simulations when the BH has reached the self-regulation and stays approximately constant.

cases, the BH bulge–mass relation would be roughly independent of redshift below $z < 6$ while the BH–galaxy mass relation would decrease smoothly with redshift. In SHDA and SLDA, the situation is more complex: initially the galaxy grows faster than the BH ($z > 3.5$), then the BH grows faster than the galaxy ($3.5 < z < 2.5$), and eventually they start tracking each other. In this case, low-mass BHs at high redshift would be considered *overmassive* with respect to their bulge, but *undermassive* if the measured quantity is the total galaxy mass, and the comparison is made to today's BH–bulge relationship (i.e. an expectation value of $\sim 10^{-3}$).

3.4 Extension to the surrounding galaxies

To check whether the impact of efficient SN explosions applies to low-mass galaxies in a general fashion, we measure the BH mass and galaxy mass relation for the other galaxies within the zoom region. We only consider galaxies that are at a distance smaller than $5\,r_{\rm vir}$ of the main galaxy. Fig. 13 shows the values of $M_{\rm BH}$–$M_{\rm gal}$ for these galaxies at various redshifts in the SLDA and SLKA runs, with the observed $M_{\rm BH}$–$M_{\rm gal}$ relation at $z = 0$ from Häring & Rix (2004). As already highlighted in the previous sections for the main galaxy, we recognize the slowing down of the growth of the most massive BH (the BH of the main galaxy) at high redshift in the run with delayed cooling SN feedback (SLDA), while

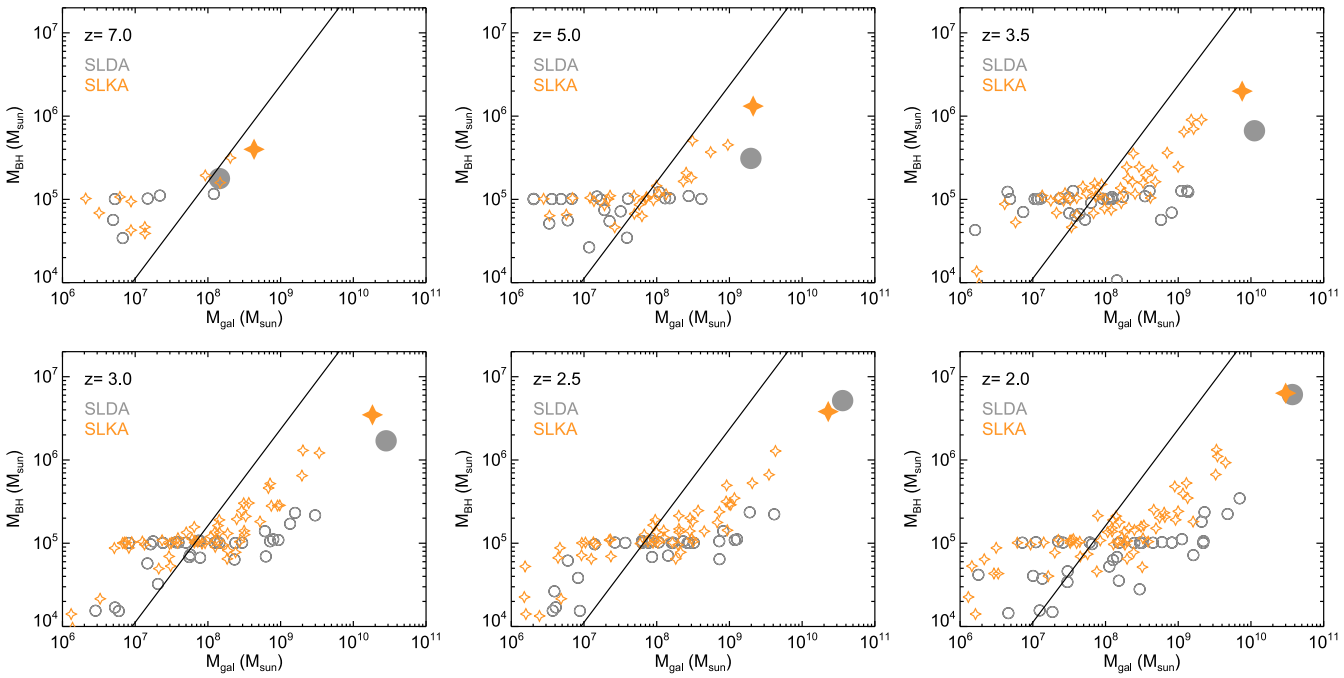

Figure 13. BH mass versus stellar mass at different redshifts as indicated in the different panels for SLDA (grey circles) and SLKA (orange stars). The main central galaxy is indicated with the filled symbol. The solid line is for the $z = 0$ observed relation from Kormendy & Ho (2013). In the SLKA simulation, BHs reach the self-regulated state very early-on, while the presence of delayed-cooling SN feedback in SLDA prevents the BHs in the less gravitationally bound bulges from growing rapidly.

the most massive BH grows quickly together with its host galaxy mass in the kinetic SN run (SLKA). Other smaller mass galaxies behave in a similar fashion: in SLKA the BH mass follows the $M_{\rm BH}$–$M_{\rm gal}$ relation (though with a larger normalization factor) and BHs start growing in galaxies of $M_{\rm gal} \simeq 10^8\,{\rm M}_\odot$; in SLDA the BH growth is delayed in low-mass galaxies, i.e. BHs start growing in more massive galaxies of $M_{\rm gal} \simeq 2 \times 10^9\,{\rm M}_\odot$, the normalization of $M_{\rm BH}$-–$M_{\rm gal}$ relation is further increased. Note that there is a plateau in our $M_{\rm BH}$–$M_{\rm gal}$ relation at low galaxy masses which is due to our choice of initial seed BH mass $10^5\,{\rm M}_\odot$ (there are BHs with smaller values because we cannot take more mass within a gas cell that is actually available.), therefore, those BHs are already overly massive for their galaxy mass.

In Fig. 14, we compare the projected gas densities of a few galaxies in SLDA and SLKA at $z = 2$. Galaxies in SLDA exhibit a porous and turbulent ISM that is driven by SN explosions. In comparison, galaxies in SLKA systematically show a more uniform gas distribution with spiral arms encompassing star-forming gas clumps but the ISM is not as porous as in the SLDA run. As discussed earlier for the main galaxy, this extra turbulence (or porosity) driven by efficient SN explosions is responsible for the inhibition of BH growth in the low-mass galaxies. Therefore, it seems plausible that this inhibition mechanism should hold on average for all galaxies with galaxy masses below a few $10^9\,{\rm M}_\odot$.

4 CONCLUSIONS

Our set of cosmological hydrodynamical simulations (the Seth suit of simulations) including either SN feedback with a kinetic model from Dubois & Teyssier (2008) or a delayed cooling prescription from Teyssier et al. (2013) show that the strength of SN feedback can dramatically alter the BH evolution at early epochs of galaxy formation, or conversely in low-mass galaxies. In the former (weak SN feedback) case, a cold gas reservoir is maintained in the centre of the galaxy sustained by inflows of dense star-forming clumps of gas, which provides an efficient gas fuelling of the central BH, driving an Eddington-limited growth early on. In the later (strong SN feedback) case, the energy release from SNe is maximized, which allows for the destruction of the dense gas clumps within the core of the galaxy. In return, the early growth of the BH is prevented due to the lack of a sustained cold gas reservoir, and this effect is systematically observed for all galaxies within the simulated region. We have shown that the characteristic SN wind velocity is of the order of $270\,{\rm km\,s^{-1}}$, a value larger than the escape velocity created by the gravitational potential of a galactic bulge of $M_{\rm b} \lesssim 10^9\,{\rm M}_\odot$ (with a fixed radius of 100 pc). This efficient mode of SN feedback prevents the rapid growth of the bulge although the disc component continues growing almost unimpeded. Once enough mass has accumulated in the centre of the galaxy, in our case driven by a major merger, the gravitational barrier prevents the removal of the cold gas and the BH can grow at a near-Eddington rate until it reaches self-regulation.

For the efficient SN feedback scenario, the bulge mass above which the accretion at the Eddington rate turns on is of $\sim$ a few $10^9\,{\rm M}_\odot$. This value is a factor of $\sim$10 smaller than the bulge mass $\sim 10^{10}\,{\rm M}_\odot$ below which abnormally light BHs are observed and exhibit a larger scatter in the BH–bulge correlation (Greene et al. 2010; Jiang et al. 2011; Graham & Scott 2015). In our simulations, only feedback from Type II SNe has been included. An extra source of energy or momentum input such as radiation pressure from young stars could enhance the wind velocities, and, thus, increase the bulge mass above which the gas ejected by stellar feedback is trapped in the bulge. Therefore, the fact that BHs undergo less growth in low-mass galaxies could be the signature of a strong stellar feedback activity, which is able to efficiently destroy the cold clumps of

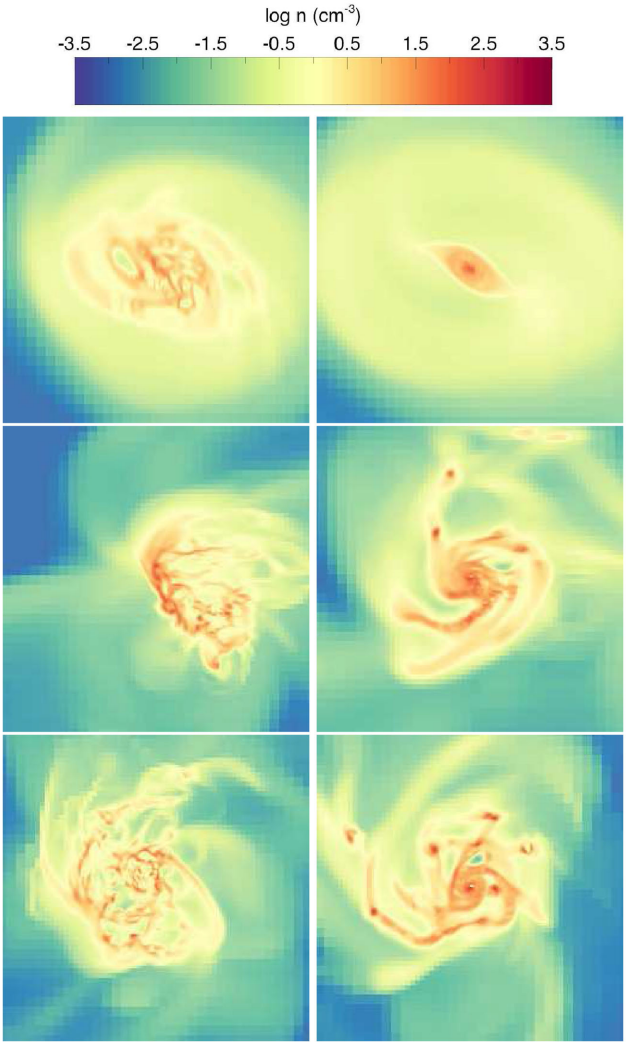

Figure 14. Projected gas densities for three other galaxies than the central one at $z = 2$ for SLDA (left column) and SLKA (right row). The images are 10 kpc across. These galaxies are not sub-structures of the main halo, and each of these images in SLDA corresponds to its equivalent in SLKA (i.e. same halo host). Galaxies simulated with delayed cooling SN feedback appear more disturbed and turbulent than galaxies simulated with kinetic SN feedback, explaining the lack of early BH growth that is a generic effect in low-mass galaxies.

star-forming gas, establishing an indirect evidence of short-lived clusters of gas below a given mass.

We have tested those two scenarios with respect to spatial resolution effects at $\Delta x = 34.8$ pc and $\Delta x = 8.7$, and our results are qualitatively independent of the adopted resolution. The Eddington growth phases appear at the same epochs: as soon as the galaxy forms for the case of kinetic SN feedback runs, and once the major merger happens in the delayed-cooling SN feedback model. Even the quantitative aspects show relatively good agreement: the bulge and disc mass are in remarkable good agreement, and the BH mass at self-regulation ends up being the same within a factor of 2.

In the Seth simulation, the 1:3 major gas-rich merger happening at $z = 3.5$ is the culprit for the triggering of the BH growth in the runs including the delayed cooling prescription for SN feedback. It does not imply that all galaxies would necessarily start growing their BH every time they endure a major merger, because they probably need to reach a mass sufficiently large enough ($M_{\rm b} > 10^9\,{\rm M}_\odot$, therefore $M_{\rm gal} > 10^9\,{\rm M}_\odot$). Once the galaxy is massive enough, a persistent bulge needs to be built, and the major merger is the leading mechanism for the Seth simulation because of the induced torques and consequent nuclear inflows. Though major mergers are the favourite candidates for creating the strongest torques, one can imagine that a massive flow of cold streams misaligned with the galactic spin would have a similar impact on the galaxy. Only extra simulations with different accretion histories, and in particular more quiescent objects, could help us to fully address that question and verify whether major mergers are systematically responsible for the ignition of BH growth in the case of efficient SN feedback.

The detailed interaction between star formation, SN feedback, BH growth and AGN feedback is probably the key to understanding the establishment of the scaling relation between BHs and their host galaxies, their redshift evolution, the overall co-evolution of BHs and galaxies, and as a consequence the properties of the galaxy population. At high redshift, estimates of BH and galaxy stellar masses of AGN samples typically find that the BH mass is larger than expected from this correlation at a given stellar mass,[4] except for submillimetre-selected galaxy (Alexander et al. 2008). Willott et al. (2010), however, based on the comparison between the BH and galaxy mass functions at $z \sim 6$, suggest that most high-z galaxies host BHs that are smaller than expected from this correlation at a given stellar mass (or host no BH at all). Similar conclusions are reached with X–ray searches of $z > 6$ galaxies: current data are sensitive to 10^5–10^6 ${\rm M}_\odot$ BHs accreting at 1–10 per cent Eddington (Willott 2011; Cowie, Barger & Hasinger 2012; Fiore et al. 2012; Treister et al. 2013; Giallongo et al. 2015; Weigel et al. 2015), but very few, if any are found, implying strong limits on the early growth of BHs in relatively small galaxies. The observational evidence, therefore, seems to suggest that high-redshift BHs in the most massive galaxies are actively accreting and are on or above the correlation with their hosts, while at lower galaxy mass the behaviour is opposite (Volonteri & Stark 2011). This new SN-regulated BH growth mechanism is a promising tool to explain why BHs in relatively low mass galaxies at high-redshift are unable to grow.

Though our method for modelling SN feedback is still very crude, it highlights two extreme regimes. Future improvements should model SNe with more physically motivated assumptions using the correct amount of momentum imparted by the explosions (Agertz et al. 2013; Kimm et al. 2015) and study the interplay between circumnuclear star formation and black hole growth. We should also consider the steady input of energy from stellar winds and radiation from bright OB associations (e.g. Hopkins et al. 2011) with dedicated simulations to probe their effect on the BH growth. Note that, even though there will be extra energy delivered by these new stellar feedback processes, it is unclear how they will interact with the gas content and star formation. As a matter of fact, stellar winds and UV radiation from young stars initiate earlier than Type II SNe, and could result in an earlier impact on the gas content of the star-forming cloud but could also limit the total amount of stars formed in one star formation cycle, thus, limiting the impact of feedback (Rosdahl et al. 2015). However, a scenario with efficient SN (or stellar) feedback mechanism is probably favoured over a less extreme mode in order to explain the low stellar mass content (e.g.

[4] High-redshift stellar masses are not equivalent to bulge masses, nor are they measured in a self-consistent way throughout different samples and at different redshift.

Hopkins et al. 2014) and the absence of bulge in dwarf galaxies (Governato et al. 2010). We dedicate these investigations to future work.

We have limited our simulations to the high-redshift ($z > 2$) environment of the progenitor of a group of galaxies and have shown that all low-mass galaxies ($M_{\rm gal} \lesssim 10^9\,{\rm M}_\odot$) within this region exhibit un-grown BHs. Though we have not explicitly demonstrated that the same mechanism is at play in the low-redshift dwarf galaxies, we argue that they should behave similarly, particularly given that galaxies at low redshift are less compact than their high-redshift equivalent, and, thus, it is easier for SN explosions to drive outflows from them. This epoch of galaxy evolution remains to be investigated and our proposed mechanism to be validated by new numerical experiments of low-redshift dwarf galaxies.

ACKNOWLEDGEMENTS

We thank the anonymous referee for useful comments, which improved the clarity of the paper. The simulations presented here were run on the DiRAC facility jointly funded by STFC, the Large Facilities Capital Fund of BIS and the University of Oxford. This research is part of the Horizon-UK project. YD and JS acknowledge support by the ERC by ERC project 267117 (DARK) hosted by Université Pierre et Marie Curie - Paris 6, and JS for support at JHU by National Science Foundation grant OIA-1124403 and by the Templeton Foundation. This work is partially supported by the Spin(e) grants ANR-13-BS05-0005 of the French Agence Nationale de la Recherche. MV acknowledges funding support from NASA, through award ATP NNX10AC84G, and from a Marie Curie Career Integration grant (PCIG10-GA-2011-303609). YD thanks Jonathan Patterson for his technical support during the course of this work.

REFERENCES

Agertz O., Kravtsov A. V., 2015, ApJ, 804, 18
Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, ApJ, 770, 25
Alexander D. M. et al., 2008, AJ, 135, 1968
Aubert D., Pichon C., Colombi S., 2004, MNRAS, 352, 376
Bardeen J. M., 1970, Nature, 226, 64
Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Bondi H., 1952, MNRAS, 112, 195
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Booth C. M., Schaye J., 2013, Scientific Reports, p. 1738
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Chapon D., Mayer L., Teyssier R., 2013, MNRAS, 429, 3114
Churazov E., Sazonov S., Sunyaev R., Forman W., Jones C., Böhringer H., 2005, MNRAS, 363, L91
Costa T., Sijacki D., Haehnelt M. G., 2014, MNRAS, 444, 2355
Cowie L. L., Barger A. J., Hasinger G., 2012, ApJ, 748, 50
Cox D. P., 1972, ApJ, 178, 159
Crain R. A. et al., 2015, MNRAS, 450, 1937
Croton D. J. et al., 2006, MNRAS, 365, 11
Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431
Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140
Dekel A., Silk J., 1986, ApJ, 303, 39
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Dubois Y., Teyssier R., 2008, A&A, 477, 79
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2010, MNRAS, 409, 985
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2012, MNRAS, 420, 2662
Dubois Y., Pichon C., Devriendt J., Silk J., Haehnelt M., Kimm T., Slyz A., 2013, MNRAS, 428, 2885
Dubois Y., Volonteri M., Silk J., Devriendt J., Slyz A., 2014a, MNRAS, 440, 2333 (Paper II)
Dubois Y., Volonteri M., Silk J., 2014b, MNRAS, 440, 1590 (Paper III)
Fabian A. C., 2012, ARA&A, 50, 455
Fiore F. et al., 2012, A&A, 537, A16
Giallongo E. et al., 2015, A&A, 578, A83
Gnedin O. Y., Zhao H., 2002, MNRAS, 333, 299
Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16
Governato F. et al., 2010, Nature, 463, 203
Graham A. W., Scott N., 2015, ApJ, 798, 54
Greene J. E. et al., 2010, ApJ, 721, 26
Haardt F., Madau P., 1996, ApJ, 461, 20
Hahn O., Abel T., 2011, MNRAS, 415, 2101
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Hu J., 2008, MNRAS, 386, 2242
Jiang Y.-F., Greene J. E., Ho L. C., 2011, ApJ, 737, L45
Kennicutt R. C., Jr, 1998, ApJ, 498, 541
Kimm T., Cen R., Devriendt J., Dubois Y., Slyz A., 2015, MNRAS, 451, 2900
King A., 2003, ApJ, 596, L27
King A. R., Lubow S. H., Ogilvie G. I., Pringle J. E., 2005, MNRAS, 363, 49
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kormendy J., Bender R., Cornell M. E., 2011, Nature, 469, 374
Krumholz M. R., Tan J. C., 2007, ApJ, 654, 304
Magorrian J. et al., 1998, AJ, 115, 2285
Mashchenko S., Couchman H. M. P., Wadsley J., 2006, Nature, 442, 539
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Newton R. D. A., Kay S. T., 2013, MNRAS, 434, 3606
Omma H., Binney J., Bryan G., Slyz A., 2004, MNRAS, 348, 1105
Ostriker E. C., 1999, ApJ, 513, 252
Peng C. Y., Impey C. D., Rix H.-W., Kochanek C. S., Keeton C. R., Falco E. E., Lehár J., McLeod B. A., 2006, ApJ, 649, 616
Planck Collaboration XVI et al., 2014, A&A, 571, A16
Pontzen A., Governato F., 2012, MNRAS, 421, 3464
Powell L. C., Slyz A., Devriendt J., 2011, MNRAS, 414, 3671
Rezzolla L., Barausse E., Dorband E. N., Pollney D., Reisswig C., Seiler J., Husa S., 2008, Phys. Rev. D, 78, 044002
Rosdahl J., Schaye J., Teyssier R., Agertz O., 2015, MNRAS, 451, 4553
Roškar R., Teyssier R., Agertz O., Wetzstein M., Moore B., 2014, MNRAS, 444, 2837
Russell H. R., McNamara B. R., Edge A. C., Hogan M. T., Main R. A., Vantyghem A. N., 2013, MNRAS, 432, 530
Sani E., Marconi A., Hunt L. K., Risaliti G., 2011, MNRAS, 413, 1479
Schramm M., Silverman J. D., 2013, ApJ, 767, 13
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Silk J., Rees M. J., 1998, A&A, 331, L1
Spaans M., Silk J., 2006, ApJ, 652, 902
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Teyssier R., 2002, A&A, 385, 337
Teyssier R., Moore B., Martizzi D., Dubois Y., Mayer L., 2011, MNRAS, 414, 195
Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, MNRAS, 429, 3068
Toro E. F., Spruce M., Speares W., 1994, Shock Waves, 4, 25
Treister E., Schawinski K., Volonteri M., Natarajan P., 2013, ApJ, 778, 130
Tremaine S. et al., 2002, ApJ, 574, 740
Tremmel M., Governato F., Volonteri M., Quinn T. R., 2015, MNRAS, 451, 1868
Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, A&A, 506, 647
Volonteri M., Stark D. P., 2011, MNRAS, 417, 2085
Wagner A. Y., Bicknell G. V., Umemura M., 2012, ApJ, 757, 136

Weigel A. K., Schawinski K., Treister E., Urry C. M., Koss M., Trakhtenbrot B., 2015, MNRAS, 448, 3167
Williams J. P., McKee C. F., 1997, ApJ, 476, 166
Willott C. J., 2011, ApJ, 742, L8
Willott C. J. et al., 2010, AJ, 140, 546
Wyithe J. S. B., Loeb A., 2003, ApJ, 595, 614

APPENDIX A: SETTING UP PARAMETERS FOR THE DELAYED COOLING MODE OF SN FEEDBACK

The time variation of the non-thermal energy component in star-forming regions follows

$$\frac{De_{\rm NT}}{Dt} = \eta_{\rm SN}\dot{\rho}_{\rm s}\frac{10^{51}\,{\rm erg}}{10\,{\rm M}_{\odot}} - \frac{e_{\rm NT}}{t_{\rm diss}} \quad (A1)$$

with the first term of the right-hand side being the injection of energy by SN explosions, and the second term being the decay rate due to the dissipation of the energy by sub-grid processes (decay of turbulence, cooling of cosmic rays, etc.). At equilibrium the two terms exactly compensate and lead to

$$e_{\rm NT} = \eta_{\rm SN}\dot{\rho}_{\rm s}\frac{10^{51}\,{\rm erg}}{10\,{\rm M}_{\odot}}t_{\rm diss}\,, \quad (A2)$$

or

$$\frac{e_{\rm NT}}{\rho} = \eta_{\rm SN}\frac{10^{51}\,{\rm erg}}{10\,{\rm M}_{\odot}}\frac{t_{\rm diss}}{t_{\rm ff}(\rho)}\epsilon_{*} \quad (A3)$$

for a Schmidt law. It comes that the corresponding non-thermal velocity dispersion is ($e_{\rm NT} = 0.5\rho\sigma_{\rm NT}^2$):

$$\sigma_{\rm NT}^2 = 2\eta_{\rm SN}\frac{10^{51}\,{\rm erg}}{10\,{\rm M}_{\odot}}\frac{t_{\rm diss}}{t_{\rm ff}(\rho)}\epsilon_{*}\,. \quad (A4)$$

The injected energy from SNe is transferred from unresolved small scales, with high gas density where it is deposited, up to resolved larger scales, with low gas density. Explicit resolved turbulence will dissipate over its own time-scale which is provided by the dynamics of shocks in the ISM. Note that, here, we derive a model for sub-grid unresolved turbulence, and the time-scale to dissipate this energy component (non-thermal energy) is turbulent crossing-time over a few resolution elements (typically one Jeans length) $l_{\rm J} = N_{\rm cell}\Delta x$ (with $N_{\rm cell} \geq 4$), so that $t_{\rm diss} \simeq l_{\rm J}/\sigma_{\rm NT}$. Therefore, equation (A4) is replaced by

$$\sigma_{\rm NT} = \left(2\eta_{\rm SN}\frac{10^{51}\,{\rm erg}}{10\,{\rm M}_{\odot}}\frac{l_{\rm J}}{t_{\rm ff}(\rho)}\epsilon_{*}\right)^{1/3}. \quad (A5)$$

We can simplify the expression further by imposing that the gas free-fall time is that at the threshold of star formation. Thus

$$t_{*}(\rho) = \frac{t_{\rm ff}(\rho)}{\epsilon_{*}} = \frac{1}{\epsilon_{*}}\sqrt{\frac{3\pi}{32\,G\rho_0}}\,. \quad (A6)$$

Finally, we can write $\sigma_{\rm NT}$ as a simple function of all known parameters for star formation and SN feedback efficiency, and resolution:

$$\sigma_{\rm NT} \simeq 48\left(\frac{\eta_{\rm SN}}{0.1}\right)^{1/3}\times\left(\frac{\epsilon_{*}}{0.01}\right)^{1/3}\times\left(\frac{N_{\rm cell}\Delta x}{4\times 10\,{\rm pc}}\right)^{1/3} \times\left(\frac{n_0}{200\,{\rm cm}^{-3}}\right)^{1/6}{\rm km\,s}^{-1}\,, \quad (A7)$$

where n_0 is the number density threshold for star formation ($n_0 = n_{\rm H,\,0}/X_{\rm H}$, $X_{\rm H} = 0.76$). The dissipative time-scale is also fully determined by the choice of $\eta_{\rm SN}$, ϵ_{*}, Δx and n_0:

$$t_{\rm diss} \simeq 0.82\left(\frac{\eta_{\rm SN}}{0.1}\right)^{-1/3}\times\left(\frac{\epsilon_{*}}{0.01}\right)^{-1/3}\times\left(\frac{N_{\rm cell}\Delta x}{4\times 10\,{\rm pc}}\right)^{2/3} \times\left(\frac{n_0}{200\,{\rm cm}^{-3}}\right)^{-1/6}{\rm Myr}\,. \quad (A8)$$

For this set of simulations we have taken $N_{\rm cell} = 4$ corresponding to the typical Jeans length at high gas densities (pressure-supported by the polytropic equation of state; see Section 2.2). Note that $\Sigma_{\rm NT}$ becomes resolution independent for our given choice of polytropic equation of state. Since $4\times\Delta x = l_{\rm J} \propto \sqrt{T_0/n_0}$, the density dependancy in equation (A7) vanishes. This simple model ensures that SN energy deposit can support the propagation of the blast wave up to $N_{\rm cell}\Delta x$ without any radiative loss since cooling is blocked up to that distance. Beyond that distance, we consider that the blast wave is resolved and radiative losses can proceed, which is ensured by the calibration of the dissipation time-scale $t_{\rm diss}$ to that of the sound-crossing time up to $N_{\rm cell}\Delta x$ for an energy deposit of $0.5\rho\sigma_{\rm NT}^2$.

APPENDIX B: GALAXY MORPHOLOGY AT LOWER RESOLUTION

In Fig. B1, we show the stellar emission in the SLDA run with lower resolution ($\Delta x = 34.8$ pc) than the SHDA run ($\Delta x = 8.7$ pc) presented in Fig. 9. The morphology of the galaxy follows qualitatively the same behaviour than the high-resolution run with the morphologically disturbed shape at $z > 3.6$, the disc-like shape with a bar component (see e.g. panels $z = 3.44$ and 3.26 in Fig. B1) during $3.1 < z < 3.8$ (see Fig. 10), the red nugget phase at $2.53 \leq z \leq 2.77$, and the formation of a spiral disc galaxy at $z \leq 2.45$ with a red bulge. As for the SHDA run, the near-Eddington BH growth starts around $z = 3.5$, once the galaxy quits the morphological disturbed phase and that the massive merger triggers the formation of a bar and nuclear inflows.

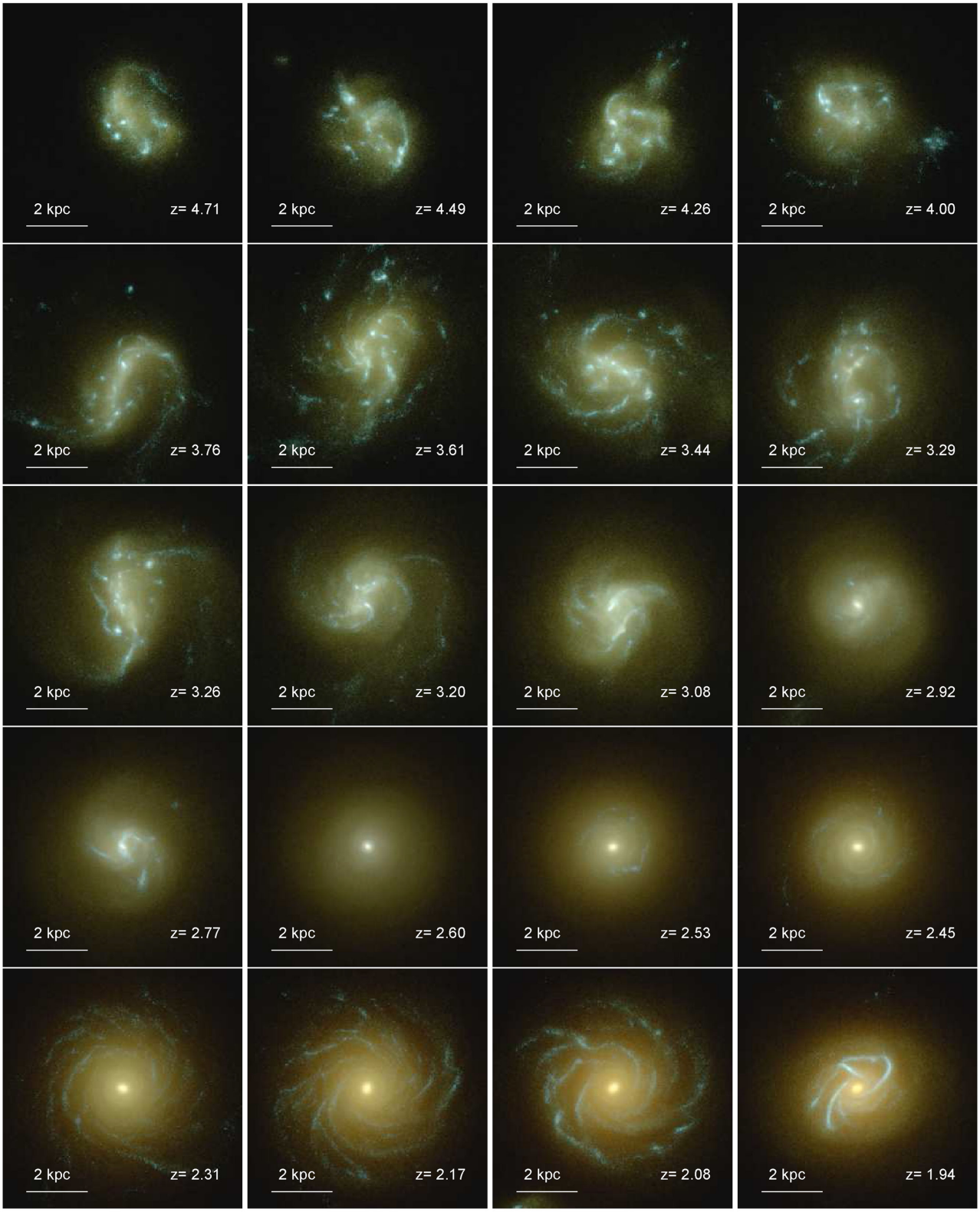

Figure B1. Same as Fig. 9 for the low-resolution run (SLDA). We have extended the stellar images to lower redshifts that we have been able to reach due to the lower resolution of the simulation. Note that though the resolution is now $\Delta x = 34.8\,\mathrm{pc}$, the galaxy looks alike the SHDA run (Fig. 9). We also see the formation of a disc spiral galaxy in the redshift range not probed by SHDA ($z \leq 2.31$).

This paper has been typeset from a TEX/LATEX file prepared by the author.